LIGHT IN THE HEART OF DARKNESS

GALAXY FORMATION FROM A NUMERICAL PERSPECTIVE

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Oscar Agertz

aus

Schweden

Promotionskomitee

Prof. Dr. Ben Moore (Vorsitz)
Prof. Dr. Daniel Wyler

Zürich, 2010

Till morbror Janne,

som aldrig hann skriva klart sin egen avhandling. Du väckte
mitt intresse för vetenskap redan innan jag kunde läsa och skriva.

Contents

Summary

This thesis addresses one of the key questions in modern astrophysics: *How do galaxies form?* In our cosmic surrounding, a great diversity of cosmic structures is observed: from tiny dwarf galaxies in the periphery of large extended disk galaxies to massive elliptical systems of stars. Today, astronomical observations have mapped, in great detail, the constituents, ages and kinematics along this sequence of diverse morphologies. However, our theoretical understanding of galaxy formation is still incomplete.

Observations of the exquisite features in the cosmic microwave background radiation have revealed details of what kind of Universe we live in. It turns out that the cosmic energy content is dominated by dark matter and dark energy, two unknown entities that are more or less essential for reconciling observations. Dark matter is five times as abundant as normal baryonic matter and clusters into extended haloes in a hierarchical fashion; small structures form first. Galaxies form in the hearts of these objects, and it is within this framework that we must be able to explain the observed multitude of galaxies.

An understanding of galaxy formation requires a model for how the hot plasma of primordial elements, existing a few hundred thousand years after the Big Bang, condenses into rapidly rotating disk galaxies of cold gas and stars. The detailed assembly of galaxies is a complex and highly non-linear process requiring the aid of numerical simulations carried out on powerful super-computers. Simulations have successfully explained the large scale features of cosmic structures such as the cosmic web of filaments. However, numerical simulations have failed to reproduce realistic disk galaxies like our own Milky Way galaxy, as they tend to be very compact, too massive and feature disk that are to small in comparison to observations. The origin of this failure is currently unknown.

In this thesis we review the theoretical framework of structure formation, from Einstein's field equations and the dynamics of space-time, to the gravitational collapse of dark matter and baryons leading to galaxy formation. We outline the numerical methods necessary for tracing the evolution of dark and baryonic matter from primordial epochs to the current time. We then focus on four issues related to numerical simulations of galaxy formation and evolution:

- We evaluate the performance of modern hydrodynamics codes relevant for galaxy formation simulations. While all methods aim to solve the equations of fluid dynamics, the way in which they are discretized, i.e. numerically implemented, differs substantially. We highlight a severe underestimation of fluid mixing in some methods and explain why and when this issue will lead to unphysical results.

- Using high-resolution simulations of isolated disk galaxies, we investigate the mechanisms responsible for driving inter-stellar turbulence. Turbulence is ubiquitous in the gas component of disk galaxies and is responsible for the stability, morphology as well as global star formation properties in all gas rich galaxies.

- We study in great detail the properties of galactic disk assembly at high redshift using hydrodynamical simulations in a cosmological context. Massive star forming galaxies at high redshifts often feature peculiar, clumpy morphologies. We demonstrate how these systems naturally can be accounted for within the current standard cosmological framework as a byproduct of epochs of intense gas accretion.

- The formation of a Milky Way like galaxy is studied in a setting similar to the previous topic. We perform a large grid of simulations aimed at demonstrating how the small scale star formation physics can significantly modify the global angular momentum content of galaxies. An inefficient formation of stars in the early Universe, i.e. ~ 10 billion years ago, motivated by less available molecular fuel for star formation in dwarfs galaxies, leads to disk galaxies that fulfill a large set of standard observational criteria. We find that if a similar effect is to be achieved via supernovae explosions, an unphysical amount of energy is required, leading to severe distortions of the gaseous disks.

Zusammenfassung

Diese Doktorarbeit beschäftigt sich mit einer der wichtigsten Fragen der modernen Astrophysik: *Wie entstehen Galaxien?* Wir beobachten in unserer kosmischen Umgebung eine grosse Vielfalt von kosmischen Strukturen; von kleinen Zwerggalaxien am Rande von grossen ausgedehnten Scheibengalaxien bis hin zu schweren elliptischen Sternsystemen. Astronomische Beobachtungen haben heutzutage sehr detailgetreu die galaktischen Komponenten, deren Alter und Kinematik entlang dieser Morphologiekette katalogisiert. Trotzdem ist unser theoretisches Verständnis noch immer unvollständig.

Beobachtungen der aussergewönlichen Merkmale der kosmischen Mikrowellenhintergrundstrahlung haben viele Details unseres Universums offengelegt. Es hat sich herausgestellt, dass der kosmische Energieinhalt unseres Universums durch dunkle Materie und dunkle Energie dominiert wird, zwei unbekannte Entitäten, die mehr oder weniger essentiell sind, um die bestehenden Beobachtungen zu erklären. Dunkle Materie ist fünf mal so häufig wie normale baryonische Materie und häuft sich in ausgedehnten Halos auf hierarchischer Art an, d.h. kleine Strukturen bilden sich zuerst. Galaxien bilden sich im Herzen dieser Objekte und dies bildet den Hintergrund vor dem wir die Vielfalt der beobachteten Galaxien erklären müssen.

Um Galaxienentstehung zu verstehen brauchen wir ein Modell der Kondensation des heissen Plasmas der Urelemente, das einige hunderttausend Jahre nach dem Urknall existierte, zu schnell rotierenden Scheibengalaxien aus kaltem Gas und Sternen. Der detaillierte Aufbau von Galaxien ist komplex und ein hochgradig nicht-linearer Prozess, der die Zuhilfenahme von numerischen Simulationen auf Superkomputern erfordert. Simulationen haben erfolgreich die grossskaligen Strukturen wie die Filamente des kosmischen Netzes erklärt. Trotzdem scheitern numerische Simulationen an der Reproduktion von realistischen Scheibengalaxien wie unsere Milchstrasse, da diese eher zu kompakt sind, zu massiv sind und zu kleine Scheiben haben im Vergleich mit den Beobachtungen. Die Ursache dieses Scheiterns ist noch nicht bekannt.

In dieser Doktorarbeit besprechen wir den theoretischen Hintergrund für Strukturentstehung angefangen von Einsteins Feldgleichungen und der Dynamik der Raum-Zeit bis zum Gravitationskollaps von dunkler und baryonischer Materie von der Frühzeit des Universums bis zur heutigen Zeit. Dann konzentrieren wir uns auf vier Fragen im Bezug auf numerische Simulationen der Galaxienentstehung und -entwicklung:

- Wir untersuchen die Leistung moderner Hydrodynamikprogramme mit Relevanz für die Galaxienentstehung. Während alle Methoden darauf abzielen die Gleichungen der Fluiddynamik zu lösen, unterscheiden sie sich beträchtlich bei der Art der Diskretisierung, d.h. ihrer numerischen Implementierung. Wir weisen auf eine deutliche Unterschätzung der Vermischung von Flüssigkeiten bei bestimmten Methoden hin und erklären warum und wann dies ein Punkt ist, der zu unphysikalischen Ergebnissen führt.

- Unter Benutzung von hochaufgelösten Simulationen von isolierten Scheibengalaxien untersuchen wir die verantwortlichen Mechanismen die interstellare Turbulenzen verursachen. Grossskalige Turbulencen sind in den Gaskomponenten von Galaxien allgegenwärtig und sind für die Stabilität, Morphologie sowie die globalen Sternentstehungseigenschaften in allen gashaltigen Galaxien verantwortlich.

- Wir studieren in grossem Detail die Eigenschaften des Scheibenaufbaus bei hoher Rotverschiebung mit hydrodynamischen Simulationen in kosmologischem Kontext. Galaxien mit hohen Sternentstehungsraten haben bei hohen Rotverschiebungen oft eigenartige, klumpige Morphologien. Wir

zeigen, wie diese Systeme natürlicherweise in unserem derzeitigen kosmologischen Standardmodel als Nebenprodukt von Epochen mit intensiver Gasakkretion entstehen.

- Die Entstehung der Milchstrasse wird in einem ähnlichen Aufbau wie oben beschrieben studiert. Wir führen ein grosses Ensemble von Simulationen durch, das darauf ausgerichtet ist, zu zeigen, wie kleinskalige Sternentstehungsprozesse signifikant den globalen Drehimpuls der Galaxien verändern können. Eine ineffiziente Sternentstehung im frühen Universum, d.h. vor ~ 10 Milliarden Jahren, motiviert durch den weniger vorhandenen molekularen Treibstoff für die Sternentstehung in Zwerggalaxien, führt zu Scheibengalaxien, die zahlreiche von üblichen Beobachtungen reproduziert. Wir kommen zu dem Ergebnis, dass, wenn ein ähnlicher Effekt mit Supernovaexplosionen erzielt werden soll, ein unphysikalisch hoher Energiebetrag von Nöten ist und dass die Gasscheiben stark verzerrt sind.

Introduction

> ‘The great spirals . . . apparently lie outside our stellar system.’
>
> — *Edwin Hubble, 1917*

In the beginning of the 20th century, Albert Einstein formulated the theory of general relativity, from which solutions existed that describe the dynamics and evolution of the Universe. At the time, the properties of the Universe outside our own Galaxy were unknown; in fact, the very *existence* of something outside was not established. Except for observations of nebulous objects by Charles Messier’s and William Herschel in the 18th century, that later turned out to be galaxies, few observations were available to test the cosmological theories[1].

The pivotal discovery made by Edwin Hubble in 1929, that all galaxies, or "island universes" as he called them, are receding from us in such a way that a linear velocity vs. distance relation exists, changed everything. This discovery was in fact theoretically predicted within the Friedmann-Lemaitre cosmology, as derived from Einstein’s equations, which pointed towards a homogenous and isotropic Universe that was expanding from an initial singular state. This became known as the Big Bang cosmology, because it predicted the cataclysmic birth of the Universe from an infinitely hot singularity, something that was a ridiculous thought to many astronomers[2].

In 1965, the *Cosmic Microwave Background* (CMB) radiation was discovered by Penzias & Wilson, hence firmly establishing cosmology as a branch of physics. The CMB radiation originates from the epoch when matter and radiation no longer could be in thermal equilibrium due to the cooling arising from the expansion of the Universe. Because the photons no longer could couple with hydrogen, they essentially moved unscattered through the Universe, being detected by us more than 13 billion years later (the age of the Universe seems to be around 13.7 billion years). This was an amazing discovery, giving us observational data to confront with theory. Being remnant heat from the Big Bang, the CMB proved to be a fundamental piece in the cosmological puzzle.

The next corner-stone of modern cosmology is the observations made by the satellite *Cosmic Background Explorer* (*COBE*) in 1992. *COBE* established that the CMB radiation had an almost perfect black-body spectrum[3] at $2.725\,\mathrm{K}$, and that the radiation was almost perfectly isotropic on the observed celestial sphere. *COBE* also probed the spatial structure of the CMB and found that it contained tiny anisotropies in the radiation temperature of the order $\Delta T/T \sim 10^{-5}$. These observations were later refined in 2003 by the *Wilkinson Microwave Anisotropy Probe* (*WMAP*), see Fig. 0.1. The inhomogeneities in the CMB sky originate from acoustic oscillations in the primordial plasma, and these tiny fluctuations are the seeds for all structure in the Universe. The spectrum of CMB fluctuations can be thought of as a fingerprint of the very young Universe, and its peaks and ripples provides us with information of what kind of Universe we live in. The model that best fit current observations is, at first glance, a peculiar one, where the energy content of our Universe is dominated by two unknown phenomena: *dark matter* and *dark energy*.

---

[1] Records show that Persian astronomers observed our closest spiral galaxy *Andromeda* as well as the *Large Magellanic Cloud* in the 10th century. Early theories of the nature of our Galaxy, and the observed nebulae in fact being "other Milky Ways" existed in the 18th century, e.g. in work by Thomas Wright and Immanuel Kant.

[2] The term *Big Bang* was coined by the British astronomer Fred Hoyle on a BBC radio show in 1949. The idea that the Universe had a beginning was philosophically troubling to Hoyle. He argued that a beginning implies a cause, and thus a creator. Being an atheist, the concept of a Big Bang was hence unacceptable.

[3] The CMB is in fact the most precisely measured black-body radiator found in nature.

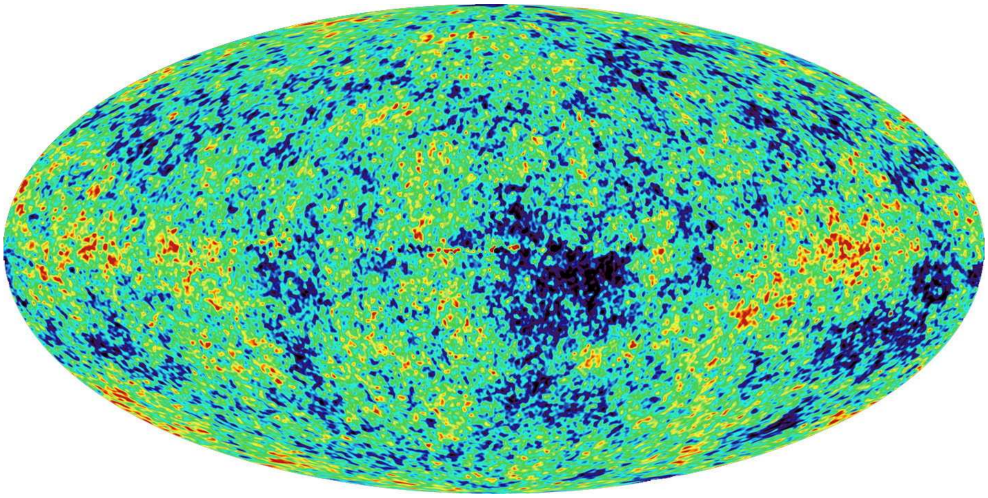

Figure 0.1: The *Cosmic Microwave Background* radiation sky as observed by the *WMAP* satellite. The light originates from the hot primordial plasma, emitted only $380\,000$ years after the Big Bang, The colours represent temperature variations of the order of $\sim 10^{-5}$ from the otherwise perfect black-body spectrum at $T = 2.725\,\mathrm{K}$. Credit: NASA/WMAP Science Team.

☐ *Dark Matter* — Dark matter is a form of matter that, unlike normal baryonic matter, does not interact strongly electromagnetically, but only via gravity. The currently favoured model of dark matter is *Cold Dark Matter* (CDM), which is thought to have non-relativistic velocities and creates structures in a bottom-up fashion (small objects form earlier than large, in a hierarchical way). Extensions of particle physics beyond the standard model, e.g. super-symmetry, can in principle accommodate this type of heavy and stable particles. Experiments such as the *Large Hadron Collider* (LHC) in CERN reach high enough energies to at least rule out some theories related to dark matter. Currently, matter constitutes only about 27 per cent of the Universe's energy content, where dark matter contributes by almost 23 per cent. This leaves only $\approx$4 per cent for normal baryonic matter.

☐ *Dark Energy* — About 73 per cent of the Universe's energy content in the unknown dark energy (also referred to as *vacuum energy*). Dark energy acts like a negative pressure at every point in space, and is required to explain the observed *accelerated* expansion of space.

This cosmological model is referred to as the ΛCDM model, where Λ denotes vacuum energy, and is the assumed model for this thesis. It should be pointed out that the notion of dark matter is not a recent concept only inferred by cosmological probes such as *COBE* and *WMAP*. In 1933, the Swiss astronomer Fritz Zwicky examined the Coma galaxy cluster and found that in order to explain the motions of the galaxies, a considerable amount of unobservable mass must reside within the cluster. His suggestion was not taken very seriously at first, but this changed almost 40 years later when the motions of stars in spiral galaxies were analyzed in detail. Based on the observed distribution of light in galaxies, Newtonian mechanics predict that the rotational velocity profile should decline at large galactic radii. In 1970, Vera Rubin made the pioneering discovery that the rotational velocities stars stay constant with radius in the outskirt of disk galaxies. This monumental discovery can be explained by an unseen distribution of matter[4], which we today refer to as *dark matter haloes*, providing the additional gravitational force to support the constant profile. In Fig. 0.2 we show a classic example of how dark matter in galaxies is inferred.

A Universe of Galaxies — A brief representation of our current understanding of the last 13.7 billion years of cosmic evolution is summarized in Fig. 0.3. Although dark matter and dark energy plays a fundamental

---

[4] Alternatively, this phenomenon can be modeled by a MOdified Newtonian Gravity (MOND), which alters the well tested $F_{\mathrm{grav}} \propto 1/r^2$ force law at large distances. In fact, this theory is supported by Vera Rubin. The theory of MOND is in general less supported by the astrophysics community compared to the dark matter scenario.

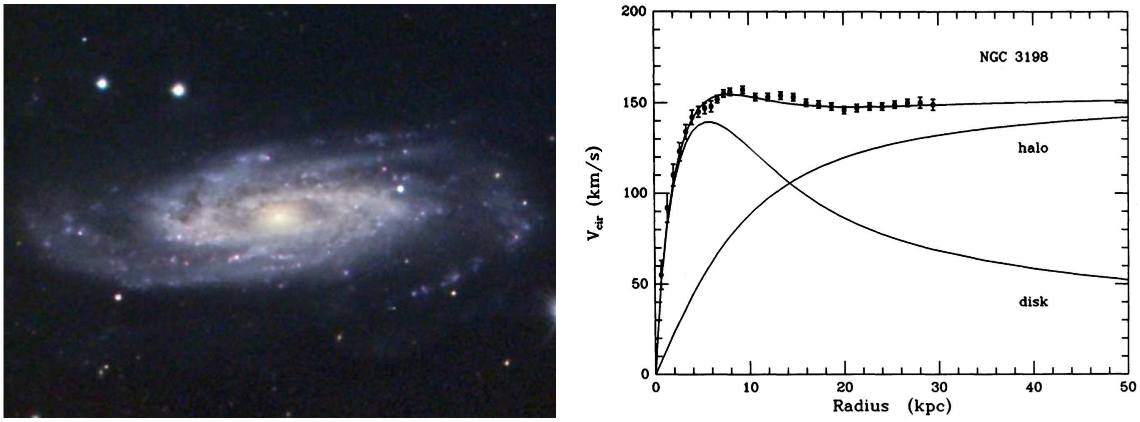

Figure 0.2: (*Left*) The Sc spiral galaxy NGC 3198 (Credit: John Vickery and Jim Matthes/Adam Block/NOAO/AURA/NSF). (*Right*) The observed rotational velocity profile (points) which flattens after $\sim 10\ \mathrm{kpc}$. This can not be explained from the observed distribution of light (the disk), but requires the existence of an unseen, extended dark matter halo (from van Albada et al., 1985).

role for the structure of the Universe, this thesis mainly concerns the fate of the baryonic component, i.e. the focus will be on the intermediate step in Fig. 0.3, modestly labelled "Development of Galaxies, Planets etc".

The current theory of structure formation, within the ΛCDM model, predicts that everything we see around us, the multitude of stars, galaxies and clusters, originated from tiny fluctuations seeded in the Big Bang. The (sole) agent for structure formation is gravity, which within our expanding Universe has led to a bottom-up structure formation scenario, in which larger and larger dark matter haloes assemble over cosmic time, and become the potential wells for the baryons to settle into. As baryons condense into the heart of dark matter haloes via gas cooling, star formation and hence galaxy formation can proceed, leading to emission of observable light. On larger scales, theory predicts that gravity should organize galaxies (and dark matter) into a network of filaments and voids; the so called *cosmic web*. Today, the cosmic web has been mapped by large scale galaxy surveys e.g. the 2dF survey, see Fig. 0.4. While the grand strokes of galaxy formation today can understood within this simple picture, many challenges remain.

It was almost a century ago that Edwin Hubble morphologically classified his observed galaxies into what today is known as *the Hubble Sequence*. Galaxies range from tiny systems with just a few thousand stars to the giant elliptical galaxies in groups and clusters each containing up to a trillion stars. Some galaxies are rotating flattened systems (spirals), others are more spherical with stars moving on random orbits (ellipticals). The governing mechanisms behind the Hubble sequence and the wide range of galaxy types are still not well understood, and complete theory of galaxy formation must be able to explain this multitude of shapes and sizes.

Today, galaxy formation is one of the main research topics within observational, theoretical and numerical cosmology. Ground and space based observatories have gathered exquisite data in multi-wavelengths with high resolution spectral information, element abundances, colour maps and kinematical data for millions of galaxies, reaching epoch when the Universe was less than ten percent of its current age[5]. Connecting our current understanding of the birth of the cosmos to the flora of galaxies and structures we see around us is one of the grand goals of modern astrophysics. In the past decades, our theoretical understanding of how galaxies form has been guided by numerical simulations. This is necessitated due to the highly non-linear nature of the involved gravitational and hydrodynamical processes. Since the first pioneering analogue[6] simulation by Erik Holmberg in 1941, who used the inverse-square law of light to mimic gravity, numerical galaxy formation

---

5. Galaxies in the early Universe feature peculiar morphologies due to violent interactions and mergers, as well as disk instabilities. Understanding the origin of this epoch in a galaxy's life will be a topic of Chap. 5 in this thesis.
6. Holmberg's simulation of interacting galaxies (using light bulbs) was confirmed 30 years later by the first simulation using a digital computer (Toomre & Toomre, 1972).

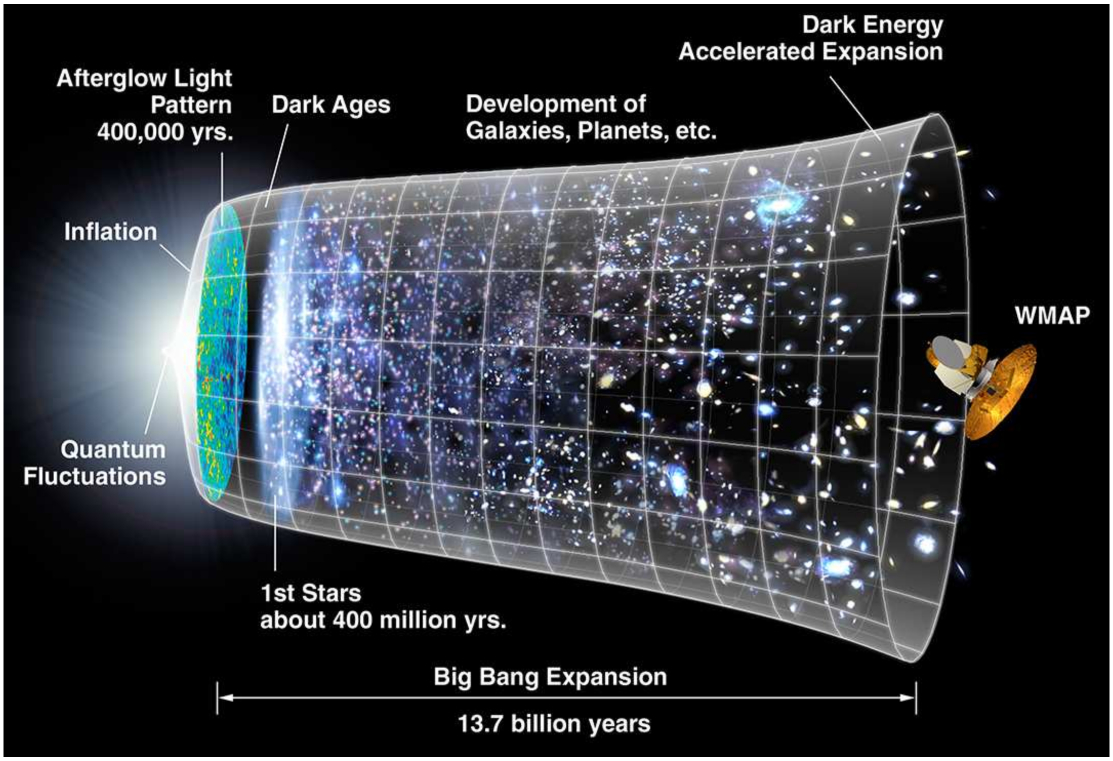

Figure 0.3: A representation of the evolution of the Universe over 13.7 billion years. The far left depicts the earliest moment we can now probe, when a period of "inflation" produced a burst of exponential growth in the Universe. For the next several billion years, the expansion of the Universe gradually slowed down due to the gravitational influence of matter. More recently, the expansion has been observed to accelerate as the repulsive effects of dark energy have come to dominate the expansion of the Universe. The afterglow light seen by *WMAP* was emitted about 380,000 years after inflation and has traversed the universe largely unimpeded since then. The conditions of earlier times are imprinted on this light; it also forms a backlight for later developments of the Universe. Credit: NASA / *WMAP* Science Team.

has made remarkable progress; fast supercomputers with tens of thousands of cores are now every day tools for theoretical astrophysicists.

In the recent decade, simulations of galaxy assembly have become increasingly sophisticated. Starting from initial conditions provided to us by the CMB observations, we are now at the stage were galaxy-like objects form in a fully cosmological setting. Modern simulations include not only gravity and hydrodynamics, but also gas cooling and heating, as well as simple models of star formation and supernovae explosions. Simulated galaxies slowly starts to look more like real galaxies, but the ultimate goal of every cosmologist, to create a realistic model of the whole Universe inside a computer, remains elusive. At the moment, the Universe is just too complicated and too large for even the fastest supercomputers. This thesis is a contribution towards this goal; to understand the formation and evolution of galaxies in a cosmological framework.

Goals and Outline of this Thesis

The primary goal of this thesis is to understand how galaxies form and evolve in a hierarchical Universe. To achieve this, we have taken a critical look at the way in which we model the astrophysical processes involved in galaxy formation. The way in which astrophysical fluids are modeled in the community can differ dramatically depending on choice of method. The same can be said about unresolved astrophysical processes such as galactic scale star formation and implementations of supernovae explosion. If these processes are contaminated with undesired numerical uncertainties and/or errors, then the predictive power of the method

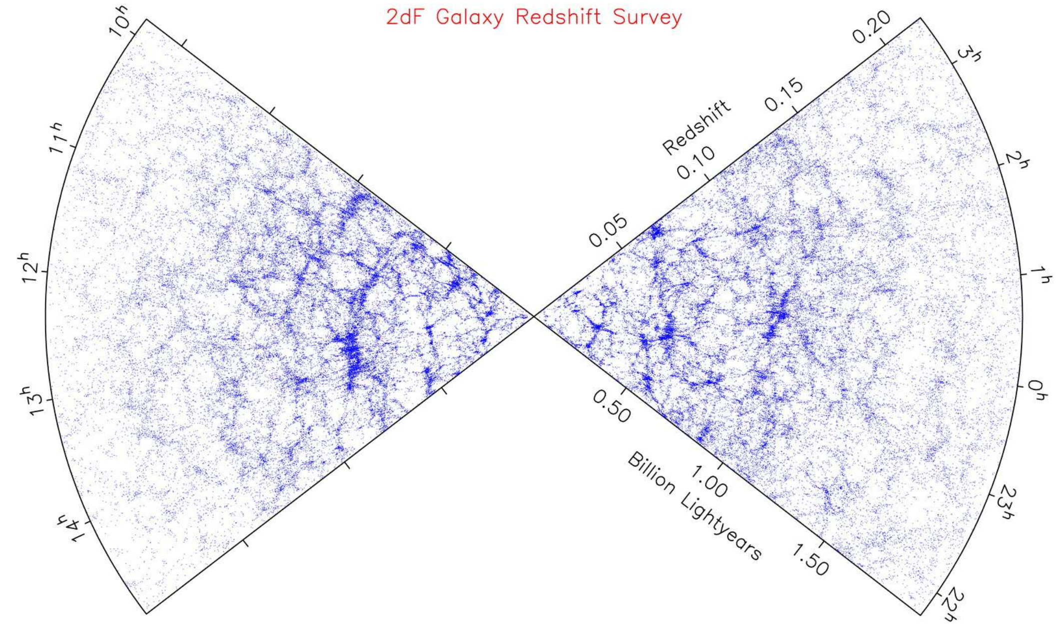

Figure 0.4: Map of the galaxy distribution produced from the completed 2dF survey. 62 559 out of the 220 929 galaxies are shown as blue dots (from the 2dFGRS website).

becomes severely limited, hence propagating erroneous conclusions. We also study, in great detail, the development and evolution of galactic turbulence, as well as the cosmological assembly of disk galaxies using state-of-the-art numerical methods.

The thesis is organized as follows:

☐ *Chapter 1: Cosmological Structure Formation* — We review the theory of structure formation, from Einstein's field equations to phenomenology of disk galaxy formation.

☐ *Chapter 2: Numerical Methods for Structure Formation* — We discuss relevant numerical methods for evolving the equations of gravity and hydrodynamics. We describe the way in which unresolved physical processes are modeled in the interstellar medium, such as star formation, supernovae type Ia and II and turbulent support.

☐ *Chapter 3: Fundamental Differences between SPH and Grid Methods* — A handful of the most commonly used fluid-dynamics codes are compared using two hydrodynamical problems: a cloud-wind interaction experiment, a.k.a *the blob test*, and a classical Kelvin-Helmholtz instability test. We demonstrate that, in contrast to grid based Eulerian methods, standard implementations of the particle based *Smoothed Particle Hydrodynamics* (SPH) scheme are unable to properly resolve multiphase interactions and fluid instabilities. We discuss the origin of this discrepancy and when this drawback will compromise physical conclusions.

☐ *Chapter 4: Large Scale Galactic Turbulence: can Self-Gravity Drive the Observed HI Velocity Dispersions?* The neutral phase of the inter-stellar medium is observed to feature super-sonic turbulent velocities. The origin of this turbulence has been debated over a long time, and the most probable driver is by many regarded to be energy injections via supernovae explosions. We demonstrate, using high resolution numerical simulations of isolated disk galaxies, that this is not the case on large, galactic scales. On these scales, gravitational instabilities in the gaseous component efficiently stirs the gas, leading to the observed levels of turbulence. Supernovae explosion will contribute to this process on smaller scales, but only above a specific star formation threshold.

☐ *Chapter 5: Disk Formation and the Origin of Clumpy Galaxies at High Redshift* — Most galaxies observed in the early Universe are remarkably different from the ones in our local Universe. The large, rapidly star forming galaxies, that are not severely distorted due to interactions, often feature extended rotating disks with supermassive clumps of stars and gas. The origin of these clumpy galaxies, and what fuels their fragmented morphology, has been a hot topic of debate in the past decade. By pushing cosmological simulations to very high resolution, we demonstrate that these galaxy morphologies are a natural outcome of disk instabilities in the early Universe, triggered by an epoch of rapid gas accretion of cold, primordial gas via cosmic gas streams.

☐ *Chapter 6: The Formation of Disk Galaxies in a ΛCDM Universe* — All attempts to form a spiral galaxy resembling our own Milky Way have been unsuccessful in the past. Simulations tend to overproduce the mass of the galactic bulge, and the disks that do form are too light and concentrated. This so called *angular momentum* problem has been given much attention in the literature, and the consensus is that supernovae explosions are responsible for shaping the galaxies into extended disks. We demonstrate that a much stronger effect on the bulge to disk ratio is achieved by considering how (in)efficiently stars form. The efficiency of star formation is believed to be severely reduced at high redshifts in assembling dwarf galaxies due to a hampered synthesis of actual star formation fuel: H_2. By crudely mimicking this process, we demonstrate that extended disk galaxies naturally can for in a ΛCDM Universe.

☐ *Chapter 7: Collaborations and the Future* — Many interesting collaborations were undertaken during the writing of this thesis. In this chapter we outline these projects and discuss the outlook on future work, both upcoming and in progress.

1.

Cosmological Structure Formation

> ‘The Universe is like a safe to which there is a combination,
> but the combination is locked up in the safe.’
>
> — *Peter de Vries*

This chapter provides an overview of the theoretical foundation for structure formation in the Universe; from general relativity to the assembly of disk galaxies.

1.1. General Relativity

Structure formation is a process fundamentally driven by gravity. It is therefore a natural starting point to review some basics of *Einstein's general theory of relativity*. The formalism presented here is, together with geometrical assumptions about our Universe, the corner stone of our standard cosmological model.

In general, an interval between two events in space-time can be written as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \qquad \text{with} \qquad \mu, \nu = 0, 1, 2, 3, \tag{1.1}$$

where the interval ds^2 is invariant under change of coordinate system. By convention, repeated suffixes imply summation where $x^0 = ct$ is the time coordinate and x^1, x^2 and x^3 are the space coordinates.

The tensor $g_{\mu\nu}$ is the *metric tensor* describing the space-time geometry. For example, the metric for flat space, i.e. *Minkowski space*, is simply $g_{\mu\nu} = \eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ rendering the interval

$$ds^2 = c^2 dt^2 - (dx^2 + dy^2 + dz^2). \tag{1.2}$$

A free particle moves according to the geodesic equation

$$\frac{d^2 x^\mu}{ds^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} = 0, \tag{1.3}$$

where Γ is *Christoffel symbol* or the *affine connection*. Written in terms of the metric tensor, it looks like

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2} g^{\mu\lambda} \left(\frac{\partial g_{\beta\lambda}}{\partial x^\alpha} + \frac{\partial g_{\alpha\lambda}}{\partial x^\beta} - \frac{\partial g_{\beta\alpha}}{\partial x^\lambda} \right). \tag{1.4}$$

From Eq. 1.3 and 1.4, it is obvious that the metric tensor plays a crucial role in general relativity; given a metric, the structure of space-time as well as the equations of motion are known.

In special relativity and Newtonian physics, conservation laws for mass, energy and momentum are fundamental. Due to the equivalence of mass and energy, these laws can in a compact way be written as

$$\frac{\partial T^{\mu\nu}}{\partial x^\nu} = 0, \tag{1.5}$$

where $T^{\mu\nu}$ is known as the *energy-momentum tensor* (or sometimes as the *stress-energy tensor*) and describes the matter distribution. In general relativity, Eq. 1.5 is incorrect because the coordinate derivative of a tensor does not necessarily yield a tensor. The correct equation is

$$T^{\mu\nu}{}_{;\nu} = 0, \tag{1.6}$$

where ‘;’ denotes the *covariant derivative* defined, for a general tensor A, as

$$A^{kl\ldots}_{pq\ldots;j} = \frac{\partial A^{kl\ldots}_{pq\ldots}}{\partial x^j} + \Gamma^k_{mj} A^{ml\ldots}_{pq\ldots} + \Gamma^l_{nj} A^{kn\ldots}_{pq\ldots} + \ldots - \Gamma^r_{pj} A^{kl\ldots}_{rq\ldots} - \Gamma^s_{qj} A^{kl\ldots}_{ps\ldots} - \ldots. \tag{1.7}$$

In the cosmologically relevant case of a perfect fluid, the energy-momentum tensor takes the form

$$T^{\mu\nu} = (\rho + p/c^2) U^\mu U^\nu - p g^{\mu\nu}, \tag{1.8}$$

where p is the pressure, ρ is the density and $U^\lambda = \mathrm{d}x^\lambda / \mathrm{d}s$ is the fluid four-velocity.

The objects introduced up to this point were available for Einstein, as he started to look for a relation between matter and the metric. The sought equations had to be linear in the second derivative of the metric to be reducible to Poisson’s equation,

$$\nabla^2 \phi = 4\pi G \rho. \tag{1.9}$$

It can be shown that the *only* tensor that is linear in second derivatives of the metric is the *Riemann tensor*

$$R^\mu{}_{\alpha\beta\gamma} = \frac{\partial \Gamma^\mu_{\alpha\gamma}}{\partial x^\beta} - \frac{\partial \Gamma^\mu_{\alpha\beta}}{\partial x^\gamma} + \Gamma^\mu_{\sigma\beta} \Gamma^\sigma_{\gamma\alpha} - \Gamma^\mu_{\sigma\gamma} \Gamma^\sigma_{\beta\alpha}. \tag{1.10}$$

This tensor was used before the work of Einstein to characterize space, and it is this tensor that gives a covariant description of curvature. The Riemann tensor is of rank four and can be contracted to render the *Ricci tensor*, and even further to the *curvature scalar*:

$$R_{\alpha\beta} = R^\mu_{\alpha\mu\beta}, \qquad R = g^{\mu\nu} R_{\mu\nu}. \tag{1.11}$$

These objects are used to form the *Einstein tensor*

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R, \tag{1.12}$$

with the important property of having, as the energy-momentum tensor, zero covariant divergence. This leads us to *Einstein’s field equations*

$$G^{\mu\nu} = -\frac{8\pi G}{c^4} T^{\mu\nu}, \tag{1.13}$$

where the numerical factors in front of the energy-momentum tensor ensures that we recover Poisson’s equation (Eq. 1.9) in the limit of weak gravitational fields. Eq. 1.13 can be modified to include terms linear in $g^{\mu\nu}$ due to the fact that $g^{\mu\nu}{}_{;\nu} = 0$ and thus $T^{\mu\nu}{}_{;\nu} = 0$ will still be valid. Modern versions of Eq. 1.13 is often written as

$$G^{\mu\nu} + \Lambda g^{\mu\nu} = -\frac{8\pi G}{c^4} T^{\mu\nu}, \tag{1.14}$$

where Λ is the so called *cosmological constant*. This term was first introduced by Einstein to construct a static Universe. Today the constant plays a central role in cosmology, describing the successfully introduced epoch of inflation as well the concept of vacuum energy. Solutions to the equations require additional constraints, which we will provide below.

1.2. The Cosmological Principle

The cosmological principle is the assertion that, on sufficiently large scales, the Universe is both homogeneous and isotropic. Homogeneity is the property of being identical at every point in space and isotropy is the property of being identical in all directions. Note that isotropy does not imply homogeneity unless one assumes that the observer is in a special place. All spherical symmetric distribution of matter would render an isotropic Universe but not necessarily a homogenous one.

The cosmological principle was introduced by Einstein and subsequent theoretical cosmologist in the early 20th century, many years before the subject of intergalactic astronomy was developed[7]. It turned out to be an extremely powerful and also very accurate fundamental principle to build the rest of cosmology upon.

---

[7] Observations and cataloging of extragalactic sources had of course been made, e.g. by the famous 18th century astronomer Charles Messier. However, the general belief was that these "nebulae" or "spiral nebulae" resided within our own galaxy. Most spiral galaxies observed in the late 19th century were in fact believed to be forming solar systems. This was the case of our neighbouring spiral galaxy, Andromeda (M31), when it was first photographed by Isaac Roberts in 1887.

Modern observations confirm that the Universe can be treated in accordance with the cosmological principle on large scales, as discussed in the introductory chapter.

From here on, we are working in units in which $c = 1$. Via geometric considerations of symmetric spaces (see e.g. Weinberg, 1972), one can show that the most general spherically symmetric homogenous space-time metric describing a Universe, incorporating the cosmological principle, is of the form

$$ds^2 = dt^2 - a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right], \tag{1.15}$$

where r, θ and ϕ are comoving spherical coordinates. This metric is often referred to as the *Friedmann metric* or, more democratically, the *Friedmann-Lamaître-Robertson-Walker metric*. Eq. 1.15 introduces two important cosmological parameters: the *expansion factor* $a(t)$, which determines the overall scale of the spatial metric, and the *curvature parameter k*, which sets the curvature of the spatial hyper-surfaces.

The expansion of space leads to a cosmological redshift of light, increasing the wavelength of observed light λ_{obs} compared to the emitted, λ_{emit}. This is conveniently parameterized using *the redshift parameter* $z \equiv (\lambda_{\mathrm{obs}} - \lambda_{\mathrm{emit}})/\lambda_{\mathrm{emit}}$, which can be related to the expansion factor as

$$1 + z(t) \equiv \frac{a(t_0)}{a(t)}, \tag{1.16}$$

where t_0 is denotes the current time and we adopt the usual definition of $a(t_0) \equiv 1$.

1.3. The Friedmann Equations

The evolution of a Friedmann Universe is obtained by solving Einstein's equations (Eq. 1.14) using the Friedmann metric (Eq. 1.15) and the stress tensor (Eq. 1.8). The $\{0, 0\}$ component and the trace of this tensor equation gives us two independent relations, the so called *Friedmann equations*. The first equation reads

$$\left(\frac{\dot{a}}{a} \right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3} \rho_{\mathrm{tot}}, \tag{1.17}$$

where ρ_{tot} is the total density with contributions from non-relativistic matter, relativistic matter, radiation and vacuum energy. The second equation is

$$\frac{2\ddot{a}}{a} + \left(\frac{\dot{a}}{a} \right)^2 + \frac{k}{a^2} = -8\pi G p, \tag{1.18}$$

where p is the pressure associated with ρ through the equation of state $p = w\rho$, w being a constant. The first Friedmann equation can be rewritten as

$$\begin{aligned} \frac{k}{a^2} &= \frac{8\pi G}{3} \rho_{\mathrm{tot}} - \frac{\dot{a}^2}{a^2} \\ &= \frac{\dot{a}^2}{a^2} \left(\frac{\rho_{\mathrm{tot}}}{(3H^2/8\pi G)} - 1 \right) \\ &= H^2(t) \left(\Omega - 1 \right). \end{aligned} \tag{1.19}$$

Here we have defined the *Hubble parameter*

$$H(t) \equiv \frac{\dot{a}}{a}, \tag{1.20}$$

which determines the rate at which the Universe is expanding, and the *density parameter*, $\Omega \equiv \rho/\rho_c$ through which the *critical density* is defined as

$$\rho_c(t) \equiv \frac{3H^2(t)}{8\pi G}. \tag{1.21}$$

From Eq. 1.19 we see that Ω determines the overall geometry of Universe. Ω refers to the total energy-matter content of the Universe and has contribution from matter (Ω_m), vacuum energy (Ω_Λ) and radiation (Ω_r). At the current epoch, $\Omega_r \ll 1$. We parametrize the first two as

$$\begin{aligned} \Omega_\mathrm{m} &\equiv \left(\frac{8\pi G}{3H_0^2}\right)\rho_0 \\ \Omega_\Lambda &\equiv \frac{\Lambda}{3H_0^2}, \end{aligned} \tag{1.22}$$

and we can also write the curvature contribution as $\Omega_k \equiv -k/a_0^2 H_0^2$. This set of density parameters, together with the current Hubble constant, defines our cosmology in the framework of the Friedmann model.

Our best estimates of these parameters comes from measurements of the Cosmic Microwave Background (CMB) fluctuations, originating from the epoch at which matter decoupled from radiation, only $\sim 380\,000$ years after the Big Bang ($z \sim 1000$). The increasingly precise measurements by *WMAP* (Komatsu et al., 2009) indicate that we live in a flat Universe ($\Omega_k = 0$) dominated by vacuum energy ($\Omega_\Lambda \approx 0.73$). Matter therefore only contributes with about 27% ($\Omega_\mathrm{m} \approx 0.27$) but of this fraction, baryonic matter contributes only by $\Omega_\mathrm{b} \approx 0.044$. This means that non-baryonic dark matter dominates the matter content of the Universe. *WMAP*, as well as observations of high-z Type Ia supernovae using the *Hubble Space Telescope* (HST) (Knop et al., 2003), suggest a value of the Hubble parameter of

$$H_0 = 100\,h\,\mathrm{km\,s^{-1}Mpc^{-1}}, \quad \text{where} \quad h \approx 0.73. \tag{1.23}$$

Eq. 1.17, together with the above definitions, can be rewritten as

$$H^2 = \frac{\dot{a}^2}{a^2} = H_0^2[\Omega_\mathrm{r}(1+z)^4 + \Omega_\mathrm{m}(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda] \tag{1.24}$$

Given an energy content, this equation tells us the dynamics of a Friedmann universe. A relation between the proper time t and z can be found via this equation and the definition of redshift (Eq. 1.16). For our $\Omega = \Omega_\mathrm{m} + \Omega_\Lambda = 1$ Universe, this reads

$$t = -\frac{1}{H_0}\int_\infty^z \frac{\mathrm{d}z}{(1+z)[\Omega_\mathrm{m}(1+z)^3 + \Omega]^{1/2}}. \tag{1.25}$$

Integrating to $z = 0$ using the *WMAP* cosmological parameters, we find that the current age of the Universe is $t_0 = 13.7\,\mathrm{Gyr}$. For a matter dominated Universe ($\Omega = \Omega_\mathrm{m} = 1$), Eq. 1.25 integrates to $t_0 = 2/3H_0 \approx 9\,\mathrm{Gyr}$, i.e. ~ 35 per cent younger than the matter dominated case. Such a young Universe is today ruled out e.g. due to the discovery of ancient stellar clusters in the halo of our galaxy, so called globular clusters, with confirmed ages surpassing 9 Gyr.

Dynamics of a Friedmann Universe — The dynamics of the Universe is highly dependent on its energy content, as demonstrated above (e.g. Eq. 1.24), and cannot be given in a closed form for an arbitrary mixture of matter, radiation and vacuum energy. Under the ansatz of $a \sim t^\beta$, a fluid equation of state ($p = w\rho$), and by taking the difference between the Friedmann equations (Eq. 1.17 and 1.18), the scale factor can be shown to evolve as

$$a(t) \sim t^{\frac{2}{3(w+1)}}, \tag{1.26}$$

where t is proper time. The equation of state of radiation is $w = 1/3$ and $w = 0$ for non-relativistic matter (sometimes called a *dust* Universe, or *Einstein-de Sitter Universe*). The latter relation reflects that pressure is negligible with respect to the mass density. These EOS imply

$$a(t) \sim \sqrt{t} \quad \text{(radiation dominated)} \tag{1.27}$$

and

$$a(t) \sim t^{2/3} \quad \text{(matter dominated).} \tag{1.28}$$

A radiation dominated Universe is a good approximation of the early Universe, which can be seen in Eq. 1.24 for $a \to 0$. If vacuum energy dominates the energy content, the first Friedmann equation simply reads

$$H^2(t) = \frac{\Lambda}{3}, \tag{1.29}$$

where we have assumed a flat Universe ($k = 0$). A trivial solution to this equation is

$$a(t) = e^{\sqrt{\frac{\Lambda}{3}}t} \quad \text{(vacuum energy dominated).} \tag{1.30}$$

This is is the so called *de Sitter Universe*[8], relevant for modeling the inflationary epoch in the early Universe. An early exponential expansion naturally solves two fundamental issues in cosmology:

☐ *The Flatness Problem:* Today's cosmological parameters appear to be fine-tuned to render a flat Universe ($k = 0$) at almost precisely the critical density ($\Omega_0 = 1$). By substituting Eq. 1.30 into Eq. 1.19, it is trivial to show that $\Omega_0 \to 1$, regardless of initial value.

☐ *The Horizon Problem:* The CMB temperature is remarkably isotropic at $T_{\rm CMB} \approx 2.725\,{\rm K}$ across the sky, with local fluctuations with amplitudes of $\Delta T \sim 10^{-5}$. The distance to the last scattering surface makes causal contact impossible across the horizon, leading to to another fine-tuning problem: *why is the radiation perfectly isotropic?*. Inflation solves the horizon problem by introducing causal contact during the first minutes after the Big Bang. The subsequent rapid expansion of space allowed for the primordial properties to "freeze in" all over the sky.

1.4. The Growth of Perturbations

The cosmological principle dictates that the Universe is homogenous and isotropic. This is of course not the case as we observe highly inhomogenous structures such as spiral galaxies and galaxy clusters. This is not in conflict with the cosmological principle as long as the Universe becomes homogenous on a certain (large) scale. Structure formation is a result of gravitational clustering, seeded by primordial inhomogeneities in the density field, which we today can observe in the CMB. The analysis of the growth of perturbations via gravity, leading to the formation of cosmic structure, is a true classic in theoretical astrophysics, dating back to the work of Jeans in the first decade of the 20th century and the classic paper by Lifshitz (1946)[9]. In this thesis, we restrict our discussion to simple non-relativistic (matter-dominated), pressure-free perturbations around the Hubble flow. Derivations for the early, radiation dominated Universe, or for super-horizon perturbations, where curvature becomes important, can be found elsewhere (e.g. Padmanabhan, 1993).

The equations of gas dynamics for a fluid in a gravitational field are:

$$\textit{Equation of continuity:} \quad \frac{{\rm d}\rho}{{\rm d}t} = -\rho \boldsymbol{\nabla} \cdot \boldsymbol{v} \tag{1.31}$$

$$\textit{Equation of motion:} \quad \frac{{\rm d}\boldsymbol{v}}{{\rm d}t} = -\frac{1}{\rho}\boldsymbol{\nabla} p - \boldsymbol{\nabla}\phi \tag{1.32}$$

$$\textit{Poisson's equation:} \quad \boldsymbol{\nabla}^2\phi = 4\pi G\rho. \tag{1.33}$$

Here, ${\rm d}/{\rm d}t$ refers to the *Lagrangian derivative*, defined as

$$\frac{\rm d}{{\rm d}t} = \frac{\partial}{\partial t} + (\boldsymbol{v} \cdot \boldsymbol{\nabla}). \tag{1.34}$$

Eq. 1.31-1.33 are in a co-moving frame, i.e. following the expansion of the Universe ($\boldsymbol{v} = \frac{\dot{a}}{a}\boldsymbol{r}$). We now considering a set of linearly perturbed variables

$$\boldsymbol{v} = \boldsymbol{v}_0 + \delta\boldsymbol{v}, \quad \rho = \rho_0 + \delta\rho, \quad p = p_0 + \delta p, \quad \phi = \phi_0 + \delta\phi, \tag{1.35}$$

where the subscript "0" denotes the average unperturbed quantities and $\delta X \ll X$. By inserting the perturbed variable in Eq. 1.31 to 1.33, it is straightforward to show that (see Appendix A for a full derivation) the evolution of the perturbed fluid equations can be characterized via the *density contrast*, $\delta \equiv \delta\rho/\rho_0$, which obeys

$$\frac{{\rm d}^2\delta}{{\rm d}t^2} + 2\frac{\dot{a}}{a}\frac{{\rm d}\delta}{{\rm d}t} = \frac{c_{\rm s}^2}{\rho_0 a^2}\nabla_{\rm c}^2\delta\rho + 4\pi G\delta\rho. \tag{1.36}$$

---

[8] Named after the Dutch mathematician Willem de Sitter. He also co-authored a paper with Albert Einstein in 1932 on cosmological models including non-light emitting matter, what we today refer to as *dark matter*, hence *Einstein-de Sitter Universe* for a matter dominated model.

[9] *On the gravitational stability of the expanding universe*, published in the Soviet Journal of Physics. This is the first work in cosmological perturbation theory to consider the scalar-vector-tensor decomposition; the decomposition of the most general linearized perturbations of the FLMR metric into components according to their transformations under spatial rotations.

Here, $\nabla_{\rm c}$ denotes differentiation with respect to the co-moving spatial coordinate. The perturbations are here assumed to be adiabatic in the sense that perturbations in pressure and density are related to the adiabatic sound speed as $\partial p/\partial\rho = c_{\rm s}^2$. In Fourier space, Eq. 1.36 reads

$$\frac{{\rm d}^2\delta}{{\rm d}t^2} + 2\frac{\dot a}{a}\frac{{\rm d}\delta}{{\rm d}t} = \delta(4\pi G\rho_0 - k^2 c_{\rm s}^2), \tag{1.37}$$

where the wave-vector k is in proper units ($\boldsymbol{k} = \boldsymbol{k}_{\rm c}/a$). This equation is a very important in astrophysical cosmology, and much can be learned by studying its solutions.

The Jeans Instability — The classical *Jeans instability*[10], studied by Jeans in 1902, is of central importance in astrophysics, especially for the process of star formation. By considering a static medium ($\dot a = 0$), and waves of the form $\delta = \delta_0 \exp i(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)$, Eq. 1.37 becomes the dispersion relation

$$\omega^2 = c_{\rm s}^2 k^2 - 4\pi G\rho_0, \tag{1.38}$$

where $\omega^2 < 0$ implies instability as usual. Depending on the sign of the right hand-side, we end up with the following two cases:

☐ $c_{\rm s}^2 k^2 > 4\pi G\rho_0$ — Pressure gradients are sufficient to prevent gravitational collapse, and the perturbation are oscillatory i.e. sound waves. Stable oscillations are found for wavelengths less than the critical *Jeans wavelength*

$$\lambda_{\rm J} = \frac{2\pi}{k_{\rm J}} = c_{\rm s}\left(\frac{\pi}{G\rho}\right)^{1/2}. \tag{1.39}$$

The mass enclosed within the sphere of diameter $\lambda_{\rm J}$, and density ρ_0, is the *Jeans mass*

$$M_{\rm J} \equiv \frac{4\pi}{3}\rho\left(\frac{1}{2}\lambda_{\rm J}\right)^3 = \frac{1}{6}\pi\rho_0\left(\frac{\pi c_{\rm s}^2}{G\rho_0}\right)^{3/2}. \tag{1.40}$$

☐ $c_{\rm s}^2 k^2 < 4\pi G\rho_0$ — The system undergoes gravitational collapse where the density contrast grows exponentially at the growth rate

$$\omega = \pm\left[4\pi G\rho_0\left(1 - \frac{\lambda_{\rm J}^2}{\lambda^2}\right)\right]^{1/2}, \tag{1.41}$$

which for very large wavelengths, $\lambda \gg \lambda_J$ becomes $\omega = (4\pi G\rho_0)^{1/2}$. In this case, the characteristic growth time for the instability becomes

$$\tau_{\rm J} = \frac{1}{\omega} = \frac{1}{(4\pi G\rho_0)^{1/2}}. \tag{1.42}$$

The Growth of Perturbations in an Expanding Universe — The second term in Eq. 1.37 modifies the classical Jeans analysis above significantly. By considering only long wavelengths, $\lambda \gg \lambda_J$, the pressure term $c_{\rm s}^2 k^2$ can be ignored, and a general solution for the density contrast can be found by solving (Heath, 1977; Carroll et al., 1992)

$$\delta(a) = \frac{\Omega_{\rm m}}{2a}\frac{{\rm d}a}{{\rm d}t}\int_0^a \frac{{\rm d}a'}{\left(\frac{{\rm d}a'}{{\rm d}t}\right)^3}. \tag{1.43}$$

This equation is best solved numerically for our current standard cosmology ($\Omega_{\rm m} + \Omega_\Lambda = 1$).

---

[10] Named after the British physicist Sir James Jeans, who considered the process of gravitational collapse within a gaseous cloud. His work was published under the title *The Stability of a Spherical Nebula* in Philosophical Transactions of the Royal Society of London in 1902.

□ *Growth in an Einstein-de Sitter Universe* ($\Omega_{\rm m} = 1, \Lambda = 0$) — In this simple case, we have already seen that $4\pi G\rho = 2/3t^2$ and $\dot{a}/a = 2/3t$. Eq. 1.37 now simplifies to

$$\frac{{\rm d}^2\delta}{{\rm d}t^2} + \frac{4}{3t}\frac{{\rm d}\delta}{{\rm d}t} - \frac{2}{3t^2}\delta = 0 \tag{1.44}$$

By seeking a power-law solution for the density contrast ($\delta = \delta_0 t^n$), the general solution is

$$\delta = \delta_+\delta_0 + \delta_-\delta_0, \tag{1.45}$$

where $\delta_+ \propto t^{2/3}$ is the growing mode and $\delta_+ \propto t^{-1}$ the (irrelevant) decaying mode. We have now found the following key result: In a matter dominated Universe, density perturbations grow as

$$\delta \propto t^{2/3} \propto a. \quad \text{(Matter dominated Universe)} \tag{1.46}$$

As it turns out, in the matter dominated phase, all modes, inside and outside the Hubble radius grow in proportion to the expansion factor. At late times, $z \lesssim 1$, vacuum energy plays a significant role and modifies this relation.

□ *Growth in the Early, Radiation Dominated Universe* — In the early Universe, the primordial perturbations are in a radiation dominated plasma, and a fully relativistic derivation, analogous to the one above, must be carried out. By using the general form of the stress tensor for a relativistic fluid (Eq. 1.8), an equation similar to Eq. 1.37 can be found, but with $4\pi G\rho \rightarrow 32\pi G\rho/3$. Here, density perturbations grow as

$$\delta \propto t \propto a^2. \quad \text{(Radiation dominated Universe).} \tag{1.47}$$

The Necessity of Collisionless Dark Matter — A fundamental outcome of Eq. 1.46 is the necessity of a dissipation-less dark matter for explaining structure formation. As described in the introductory chapter, the inferred density fluctuations in the CMB are on the order of $\delta \sim 10^{-5}$. This is at $z_{\rm rec} \sim 1000$, meaning that a linear growth would only lead to $\delta \sim 10^{-2}$ at $z = 0$. Hence, in a Universe dominated by baryons, gravitational clustering is unable to explain the existence of the collapsed structures, as structure formation would require fluctuations larger than $\delta \sim 10^{-3}$ at the epoch of recombination.

From the epoch of matter-radiation equality ($z \approx 2 \times 10^4$), to the epoch recombination, perturbations of mass less than $M_{\rm J} = 3.75 \times 10^{15}/(\Omega_{\rm bar}h^2)^2$ were not unstable but oscillated as sound-waves, the so called *baryonic acoustic oscillaitons*. A collisionless dark matter component is unaffected by this, and grows freely through the pre-recombination epoch. Post recombination, baryons will, crudely speaking, fall into the far more advanced dark matter perturbations, leading to galaxy formation.

The Zel'Dovich Approximation — A simple approximation to describe the onset of non-linear gravitational evolution was developed by Zel'Dovich (1970). The *Zel'Dovich approximation* elegantly reveals the nature of the structures that initially form in the Universe, and has become a valuable tool for setting initial conditions for numerical simulations of cosmological structure formation.

At first, we assume that all perturbations are in the linear limit and that we are allowed to write the trajectory of a particle in an Eulerian fashion as

$$\boldsymbol{x}(a, \boldsymbol{q}) = \boldsymbol{q} + \boldsymbol{L}(a, \boldsymbol{q}) \tag{1.48}$$

where $\boldsymbol{q}$ is the Lagrangian coordinate of the original position and $\boldsymbol{L}(a, \boldsymbol{q})$ is the displacement. The density contrast for this trajectory can be written as

$$\delta(a, \boldsymbol{k}) = \int d^3\boldsymbol{x} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\delta(a, \boldsymbol{x}) = \int d^3\boldsymbol{q} e^{-i\boldsymbol{k}\cdot\boldsymbol{q} - i\boldsymbol{k}\cdot\boldsymbol{L}(a,\boldsymbol{q})} - (2\pi)^3\delta_D(\boldsymbol{k}). \tag{1.49}$$

The particles are expected to move slowly in the linear regime. This makes it possible to expand Eq. 1.49 in a Taylor series in $\boldsymbol{k} \cdot \boldsymbol{L}$. To first order, we find

$$\delta(a, \boldsymbol{k}) \cong -\int d^3\boldsymbol{q} e^{-i\boldsymbol{k}\cdot\boldsymbol{q}}[i\boldsymbol{k} \cdot \boldsymbol{L}(a, \boldsymbol{q})] = -\int d^3\boldsymbol{q} e^{-i\boldsymbol{k}\cdot\boldsymbol{q}}(\nabla_{\boldsymbol{q}} \cdot \boldsymbol{L}). \tag{1.50}$$

This means that $\delta(a, \boldsymbol{k})$ is, to first order, the Fourier transform of $-\nabla_q \cdot \boldsymbol{L}$. In real space, this implies that

$$\nabla \cdot \boldsymbol{L}(a, \boldsymbol{q}) = -\delta(a, \boldsymbol{k}) = -\frac{2}{3}H_0^{-2}a\nabla \cdot (\nabla\phi) \tag{1.51}$$

where the last equality is valid for a $\Omega = 1$ Universe. This means that particle displacements that are consistent with a growing mode $\delta(a) \propto a$ are given by

$$\boldsymbol{L}(a, \boldsymbol{q}) = -a(\nabla\psi) \equiv a\boldsymbol{u}(\boldsymbol{q}), \quad \text{where} \quad \psi \equiv \frac{2}{3}H_0^{-2}\phi. \tag{1.52}$$

This relation leads us to the formulation of the Zel'Dovich approximation: The proper Eulerian position $\boldsymbol{r}$ of a particle is related to its Lagrangian position $\boldsymbol{q}$ by

$$\boldsymbol{r}(t) \equiv a(t)\boldsymbol{x}(t) = a(t)[\boldsymbol{q} + a(t)\boldsymbol{u}(\boldsymbol{q})], \tag{1.53}$$

where $\boldsymbol{x}(t)$ is the comoving Eulerian coordinate. The Zel'Dovich approximation is not valid in the strong non-linear regime but can successfully be used in the quasi-linear regime. It allows us to investigate what kind of structures that are the first to form in the following way. Conservation of mass tells us that the perturbed density, $\rho(\boldsymbol{r}, t)$, is related to the initial unperturbed density, $\overline{\rho}$, as

$$\rho(\boldsymbol{r}, t)d^3\boldsymbol{r} = \overline{\rho}d^3\boldsymbol{q}. \tag{1.54}$$

It follows that

$$\rho(\boldsymbol{r}, t) = \frac{\overline{\rho}/a^3}{\det(\partial x_j/\partial x_i)} = \frac{\rho_{\rm b}(t)}{\det[\delta_{ij} + a(t)[\partial u_j/\partial x_i]} \tag{1.55}$$

where $\rho_{\rm b} = \overline{\rho}/a^3$. The denominator is a determinant of a real, symmetric matrix. By diagonalizing it and writing the eigenvalus of $(\partial u_j/\partial x_i)$ as $-\lambda_1(\boldsymbol{q})$,$-\lambda_2(\boldsymbol{q})$ and $-\lambda_3(\boldsymbol{q})$, the perturbed density is given by

$$\rho(\boldsymbol{r}, t) = \frac{\rho_{\rm b}(t)}{[1 - a(t)\lambda_1(\boldsymbol{q})][1 - a(t)\lambda_2(\boldsymbol{q})][1 - a(t)\lambda_3(\boldsymbol{q})]} \tag{1.56}$$

Let us assume that the eigenvalues are different and can be ordered as $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Because $a(t)$ grows, there will be a particular moment, and coordinate $\boldsymbol{q}$, for which $[1 - a(t)\lambda_1(\boldsymbol{q})] = 0$. This singularity of Eq. 1.56 corresponds to the formation of a pancake (sheet-like structure) by contraction along one of the principal axes. These are the first structures to be expected in gravitational clustering. Subsequent collapse along the other axes are then expected to lead filaments and virialized dark matter haloes, see discussion below.

The Zel'Dovich approximation is a key ingredient when generating initial conditions for cosmological $N-$body simulations (see Chap. 2.2). By placing a set of particles on a uniform grid, Eq. 1.48 can be used to displace particles according to a well defined spectrum of fluctuations (see Chap. 1.6). Hence, by knowing the initial power spectrum $P(k)$ we can appropriately sample $\delta_k \sim \sqrt{P(k)}$ and reconstruct a field of displaced particles that in turn can be integrated by $N-$body methods.

1.5. Spherical Collapse, Virialization and the Formation of Dark Matter Haloes

While the derivations in the previous sections are useful for characterizing the evolution of the density field in the linear regime ($\delta \ll 1$), it fails to describe the evolution of locally self-gravitating structures. The non-linear collapse of over-dense regions and their equilibria is a topic best explored using numerical N-body simulations. However, an approximate treatment can be carried out using a model known as *the spherical top-hat* model.

A spherical perturbation behaves dynamically like an embedded Universe of slightly higher density compared to the otherwise uniform Universe. This means that the Friedmann equations in Eq. 1.17 and 1.18 can be used to study the evolution. For a matter dominated world model, the Friedmann equations can be combined into

$$\dot{a}^2 = H_0^2\left[\Omega_{\rm m,0}\left(\frac{1}{a} - 1\right) + 1\right]. \tag{1.57}$$

For a slightly over-dense region ($\Omega_{\mathrm{m},0} > 1$), the region behaves like a closed Universe destined to collapse. The general solution to Eq. 1.57 can be written in parametric form[11] as

$$\begin{aligned} R &= A(1-\cos\theta) \qquad & t &= B(\theta - \sin\theta) \\ A &= R_0 \frac{\Omega_{\mathrm{m},0}}{2(\Omega_{\mathrm{m},0}-1)} \qquad & B &= \frac{\Omega_{\mathrm{m},0}}{H_0 2(\Omega_{\mathrm{m},0}-1)^{3/2}}. \end{aligned} \tag{1.58}$$

The perturbation will reach its maximum size for $\theta = \pi$, the so called *turnaround radius*;

$$R_{\max} = 2A = R_0 \frac{\Omega_{\mathrm{m},0}}{\Omega_{\mathrm{m},0}-1} \quad \text{at } t = t_{\max} = \pi B, \tag{1.59}$$

and complete collapse have occurred at $\theta = 2\pi$.

By expanding $\cos\theta$ and $\sin\theta$ to fifth order, the equation for R becomes

$$\frac{R(t)}{R_{\max}} \approx \frac{1}{4}\left(6\pi\frac{t}{t_{\max}}\right)^{2/3}\left[1 - \frac{1}{20}\left(6\pi\frac{t}{t_{\max}}\right)^{2/3}\right]. \tag{1.60}$$

The term outside of the square brackets is just the expansion of the background Einstein-de Sitter Universe (c.f. Eq. 1.46). The terms within the square bracket is therefore the linear theory expression for the growth of a perturbation. We can now express the evolution of the linear over-density as

$$\begin{aligned} \delta_{\mathrm{lin}}(t) = \frac{\rho}{\rho_0} - 1 = \left(\frac{a}{R}\right)^3 - 1 \quad &= \quad \left[1 - \frac{1}{20}\left(6\pi\frac{t}{t_{\max}}\right)^{2/3}\right]^{-3} - 1 \\ &\approx \quad \frac{3}{20}\left(6\pi\frac{t}{t_{\max}}\right)^{2/3}. \end{aligned} \tag{1.61}$$

This tells us that, in linear theory, when a perturbation has reached turnaround ($t = t_{\max}$) and breaks away form the background, it has an amplitude of

$$\delta_{\mathrm{lin}}^{\mathrm{turn}} = \frac{3}{20}(6\pi)^{2/3} = 1.06. \tag{1.62}$$

At full collapse ($t = 2t_{\max}$) the amplitude is

$$\delta_{\mathrm{lin}}^{\mathrm{turn}} = \frac{3}{20}(12\pi)^{2/3} = 1.686. \tag{1.63}$$

The fact that overdensities are expected to be collapsed at $\delta \sim 1.7$ in linear analysis is important for the analytical treatment of halo mass functions in Sect. 1.7. The actual nonlinear density contrast at turnaround is

$$\frac{\rho_{\max}}{\rho_0} = \left(\frac{a(t_{\max)}}{R_{\max}}\right)^3 = \left[\frac{1}{4}\left(6\pi\frac{t}{t_{\max}}\right)^{2/3}\right]^3 = \frac{9\pi^2}{16} \approx 5.55. \tag{1.64}$$

Thus, at the time of turnaround, the density of the sphere is already 5.55 times greater than the background.

Virialization — What happens at the time of collapse, $t = 2t_{\max}$? While a gaseous system would build up pressure gradients to break the collapse, the end-state of a perfectly spherical, collisionless system is in principle a black-hole, as the collapsing system would eventually end up in an infinitesimal volume, regardless of mass. In reality, inhomogeneities and environmental torques are expected to modify this catastrophic fate, leading to a redistribution of kinetic energy into random motions i.e. *virialization*. The dark matter structures resulting from this process are called *haloes*. This is process is principle a problem best studied via direct numerical integration of $N-$body systems, which will be discussed in Chap. 2, but some important analytical relations can be derived.

The virial theorem states that the gravitational potential energy of a self-gravitating system in equilibrium amounts to twice its negative internal kinetic energy. i.e.

$$U = -2K. \tag{1.65}$$

---

[11] Note that this solution is analogous to the falling pole problem in mechanics.

At turnaround, there is no kinetic energy ($K_{\rm turn} = 0$). From conservation of energy and the virial theorem, the gravitational potential energy at turnaround is

$$U_{\rm turn} = U_{\rm vir} + K_{\rm vir} = \frac{1}{2} U_{\rm vir}. \tag{1.66}$$

Assuming a uniform sphere, $U_{\rm turn} = -3GM^2/5R_{\rm max}$, we find that $R_{\rm vir} = 1/2R_{\rm max}$; *the final virialized object has collapsed by a factor of two*. The virialized overdensity is then (c.f. Eq. 1.64)

$$\frac{\rho_{\rm vir}}{\rho_0} = \left(\frac{a(2t_{\rm max})}{R_{\rm max}/2}\right)^3 = 18\pi^2 \approx 178 \tag{1.67}$$

Although this is highly idealized situation, numerical simulations indicate that this result is roughly correct. In a full ΛCDM cosmology, the virial overdensity is no longer a constant in time. Bryan & Norman (1998) provided a simple fitting formula for this case;

$$\Delta_c = \frac{\rho_{\rm vir}}{\rho_{\rm crit}} = 18\pi^2 + 82x - 39x^2, \quad (\Omega_m + \Omega_\Lambda = 1), \tag{1.68}$$

where $x = (1 + a^3\Omega_{\Lambda,0}/\Omega_{\rm m,0})^{-1} - 1$. Note that Δ_c is relative to the *critical* density, not the background. In the current concordance cosmology described in Sect. 1.3 ($\Omega_{\rm m,0} = 0.27$, $\Omega_\Lambda = 0.73$), the virial overdensity at $z = 0$ becomes $\rho_{\rm vir}/\rho_0 \approx 360$, and $\Delta_c \approx 97$.

The Structure of Cold Dark Matter Haloes — The equilibrium structure of dark matter haloes has been a hot topic for the last decades, and has mostly been driven by the advancement in numerical simulations of $N-$body systems. An isothermal gas, or a collisionless system at a constant velocity dispersion σ can analytically be shown[12] to follow a *singular isothermal sphere*

$$\rho_{\rm iso} = \frac{\sigma^2}{2\pi G r^2}. \tag{1.69}$$

The profile is unphysical as it has infinite density at $r = 0$, and $M(r) \sim r$, i.e. it has infinite mass as $r \rightarrow \infty$. This is often remedied by introducing a central core or by modifying the distribution function (King modeling). However, due to the simplicity of the isothermal profile, it is still a widely used for e.g. fitting cluster gas profiles.

A pivotal study of cold dark matter halo density profiles for different masses and cosmological models was carried out by Navarro et al. (1996). This work indicated that a general density profile, often referred to as the *NFW-profile*, could fit all simulated haloes. We will refer to this profile throughout this thesis, and will therefor define it in detail. The density profile is

$$\rho(r) = \frac{\rho_{\rm c}\delta_c}{(r/r_s)(1 + r/r_s)^2}, \tag{1.70}$$

where $r_{\rm s}$ is the scale radius, $\rho_{\rm c}$ is given by Eq. 1.21, and δ_c is a dimensionless density given by

$$\delta_c = \frac{\Delta_{\rm c}}{3} \frac{c^3}{[\ln(1 + c)] - c/(1 + c)]}, \tag{1.71}$$

where c is the *halo concentration*. The scale radius corresponds to the radius where

$$\left.\frac{d\log\rho(r)}{d\log r}\right|_{r=r_{\rm s}} = -2. \tag{1.72}$$

The standard usage is to set $\Delta_{\rm c} = 200$, hence defining the virial mass M_{200} as the mass within the virial radius r_{200}, defined as the radius of a sphere with a mean interior density $200\rho_{\rm c}$. The halo structure is hence a two parameter function, completely determined by M_{200} (or equivalently r_{200} or V_{200}) and the concentration parameter (or equivalently δ_c). Note that the scale radius is related to the virial radius as $r_{\rm s} = r_{200}/c$. The slope of the NFW-profile approaches -1 at small radii and -3 at large.

While the functional behaviour in the outer parts of CDM haloes are pretty well established, the correct slope in the inner parts is still under debate and many functional forms exists in the literature (for a recent overview,

---

[12] Assume a Maxwell-Boltzmann distribution function and solve for the equation of hydrostatic equilibrium and Poisson's equation, see e.g. Binney & Tremaine (2008)

see Stadel et al., 2009). Recent work suggest a more general profile; the *Einasto*-profile (Einasto, 1969) to be the best fit to CDM haloes. It is important to note that these results are derived from purely dissipation-less simulations, and as the inner structure is believed to change due to the process of galaxy formation.

1.6. The Spectrum of Fluctuations

As long as $\delta(\boldsymbol{x},t) \ll 1$, each mode $\delta_{\boldsymbol{k}}(t)$ of the dark matter will evolve independently and in proportion to the expansion factor. A convenient way of denoting the cosmic evolution of the density contrast is

$$\delta_{\boldsymbol{k}}(t) = T_{\boldsymbol{k}}(t,t_{\mathrm{i}})\delta_{\boldsymbol{k}}(t_{\mathrm{i}}), \tag{1.73}$$

where $T_{\boldsymbol{k}}(t,t_{\mathrm{i}})$ is called the *transfer function* and t_{i} is the initial time. The behaviour of all modes depends on the expansion parameter a, the size of a at matter-radiation equilibrium a_{eq}, and the time t_{enter} the mode enters the horizon. Because of this, the transfer function is fairly straightforward to calculate in the linear regime.

We have, to this point, seen that in order to determine $\delta(\boldsymbol{x},t)$ or $\delta_{\boldsymbol{k}}(t)$ at a time t, the exact $\boldsymbol{x}$ or $\boldsymbol{k}$-dependance at some initial time t_{i} must be known. This way of treating things are often very cumbersome and one is not always interested in this exact form of $\delta(\boldsymbol{x},t)$ but rather in its statistical properties. Let each Fourier mode $\delta_{\boldsymbol{k}}(t_{\mathrm{i}})$ be a Gaussian random variable. This implies that

$$\langle \delta_{\boldsymbol{k}}(t_{\mathrm{i}})\delta^*_{\boldsymbol{p}}(t_{\mathrm{i}})\rangle = (2\pi)^3 P(\boldsymbol{k},t_{\mathrm{i}})\delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{p}), \tag{1.74}$$

where $P(\boldsymbol{k},t_{\mathrm{i}})$ is the power spectrum of $\delta(\boldsymbol{x},t_{\mathrm{i}})$, $\langle\cdots\rangle$ denotes an ensemble average and $*$ denotes the complex conjugate. Using Eq. (1.73) we find that

$$\begin{aligned}\langle \delta_{\boldsymbol{k}}(t)\delta^*_{\boldsymbol{p}}(t)\rangle &= T_{\boldsymbol{k}}(t,t_{\mathrm{i}})T^*_{\boldsymbol{p}}(t,t_{\mathrm{i}})\langle \delta_{\boldsymbol{k}}(t_{\mathrm{i}})\delta^*_{\boldsymbol{p}}(t_{\mathrm{i}})\rangle \\ &= (2\pi)^3 |T_{\boldsymbol{k}}(t,t_{\mathrm{i}})|^2 P(\boldsymbol{k},t_{\mathrm{i}})\delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{p}),\end{aligned} \tag{1.75}$$

which means that the evolution of the power spectrum is given by

$$P(\boldsymbol{k},t) = |T_{\boldsymbol{k}}(t,t_{\mathrm{i}})|^2 P(\boldsymbol{k},t_{\mathrm{i}}). \tag{1.76}$$

In the cosmological context, $P(k)$ describes the strength of density modes associated with the wave vector $\boldsymbol{k}$. For example, if $P(k) \sim k$ there is more power for large values of k i.e. more power on small scales.

The power spectrum can be used to study the statistical evolution of any random field but it is only in the case of Gaussian random fields that the spectrum contains *all* information. This is due to the following fact: In previous sections, we only had to consider the amplitude of density fluctuations. Generally speaking, $\delta_{\boldsymbol{k}}$ is a complex variable that can be decomposed into two real variables, the amplitude $D_{\boldsymbol{k}}$ and the phase $\phi_{\boldsymbol{k}}$ as follows

$$\delta_{\boldsymbol{k}} \equiv D_{\boldsymbol{k}} \exp(i\phi_{\boldsymbol{k}}). \tag{1.77}$$

The power spectrum contains no information about the phase, only the amplitude. When $\delta_{\boldsymbol{k}}$ is Gaussian distributed, the phases are uniformly distributed and contain no information. In this case, $P(k)$ contains all information about the system. If the $\delta_{\boldsymbol{k}}$s are non-Gaussian, Eq. (1.75) would carry additional factors. As it turns out, theories of the inflationary epoch predict that the statistics of initial perturbations are in fact Gaussian. It is therefore safe to say that the statistics in the linear regime is preserved. In the non-linear regime fluctuations grow larger than unity, modes start to couple and the Gaussian assumption can not be guaranteed.

It is common to assume that the initial power spectrum is a power law. For an isotropic and homogenous Universe the relation is simply

$$P(k) = |\delta_k|^2 \propto k^n. \tag{1.78}$$

This assumption is strongly supported by most theories of inflation. It is often more revealing to study the power in each logarithmic interval of k. The quantity of interest is then

$$\Delta^2_k \equiv \frac{k^3 P(k)}{2\pi^2} \tag{1.79}$$

and from simple arguments (e.g. Padmanabhan 2002), the evolution before and after matter-radiation equilibrium is

$$\Delta_k^2 = \begin{cases} L^2(k)k^{n-1}(a/a_{\rm eq})^2 & \text{for} \quad k_{\rm eq} < k \\ k^{n+3}(a/a_{\rm eq})^2 & \text{for} \quad k < k_{\rm eq}, \end{cases}, \tag{1.80}$$

where $L \simeq 5\ln(\lambda_{\rm eq}/\lambda)$ describes the growth of a scale λ between $t_{\rm eq}$ and $t_{\rm enter}$ through the relation $\delta_\lambda(t_{\rm eq}) = L\delta_\lambda(t_{\rm enter})$.

The index $n = 1$ is of special interest. First of all, this value of n is predicted by several inflationary models and seems to be the value we observe in the Universe (Komatsu et al., 2009). Also, $\Delta_k^2(t_{\rm enter})$ will be independent of k and all scales enter the Hubble radius *with the same amplitude*. In the range $k_{\rm eq} < k$, all scales will have almost the same power, except for the weak factor $L^2(k)$ giving smaller scales slightly more power than the larger scales.

The above discussion of the power spectrum is approximate. Eq. (1.80) only roughly describes how power is transferred. The transfer function is in principle calculated by solving the coupled Boltzmann equations for all relevant species of particles, and several packages are available in the literature, e.g. COSMICS (Bertschinger, 1995) and CMBFAST (Seljak & Zaldarriaga, 1996). The solutions cannot be expressed in closed form and one is forced to use numerically fitted versions (Eisenstein & Hu, 1999). This treatment provides us with the shape of the evolved spectrum, but not the normalization which requires observational input. This is usually done by normalizing it to be in accordance with the observed root-mean-squared amplitude of galaxy fluctuations within $8\,h^{-1}$Mpc, σ_8. Current best-fit measurements (Komatsu et al., 2009) indicate that $\sigma_8 = 0.812 \pm 0.026$

1.7. The Abundance of Collapsed Objects

Equipped with a theory of how primordial perturbations grow, and the tools to statistically describe their scale dependent clustering strength, we can derive a general formula for the co-moving number density of virialized objects. Press & Schechter (1974) realized that haloes could be associated with peaks in the Gaussian random density field of dark matter in the early universe. Using the relatively simple statistics of Gaussian random fields, in which higher statistics of higher order than the power spectrum $P(k)$ are irrelevant, they were able to derive the following form for the distribution of dark matter halo masses such that the number of halos per unit volume in the mass range M to $M + dM$ is

$$n(M,z)dM = \frac{\rho_0}{M}\frac{d\nu(M,z)}{dM}f(\nu)dM, \tag{1.81}$$

where ρ_0 is the mean density of the Universe at $z = 0$ and $f(\nu)$ is a mass function. Also,

$$\nu(M,z) = \frac{\delta_c(z)}{\sigma(M,z)}, \tag{1.82}$$

where δ_c is the critical overdensity required for a collapse of a spherical region at redshift z ($\delta_c = 1.68$ for an Einstein-de Sitter Universe, c.f. Sect. 1.5) and σ is the r.m.s. linear over-density in a sphere containing the mass M, which for a Gaussian density field can be calculated from

$$\sigma^2(R) = \frac{1}{2\pi^2}\int_0^\infty P(k)W^2(kR)k^2dk. \tag{1.83}$$

Here the radius R is connected to the mass of an initial spherical object through the usual $M = \frac{4\pi}{3}\rho_0 R^3$. The window function W is a Fourier transformed spherical top-hat filter of the form

$$W(x) = 3(\sin x - x\cos x)/x^3. \tag{1.84}$$

The most well known form of $f(\nu)$ is the classic *Press-Schechter* (PS) mass function of the form

$$f(\nu) = \sqrt{\frac{2}{\pi}}\exp(-\nu^2/2). \tag{1.85}$$

Press & Schechter were well aware of the limitations of a formalism that basically only studies spherical over-densities from linear theory, but it has turned out to be very successful in general predictions of cluster abundances in both observations and simulations. However, the PS formalism seems to predict too many low-mass clusters and too few high-mass clusters. More elaborate mass function have been proposed by (e.g.

Sheth & Tormen, 1999; Reed et al., 2007; Tinker et al., 2008). Accurate matching to mass functions from fully non-linear $N-$body simulation is achieved by tuning unknown parameters to the simulations.

1.8. Galaxy Formation

This is the main topic of this thesis, which will be comprehensively studied in Chap. 5 and 6. Our prevailing picture of galaxy formation emerged more than 30 years ago (White & Rees, 1978; Fall & Efstathiou, 1980). Within the framework of the Λ Cold Dark Matter (ΛCDM) scenario (Komatsu et al., 2009), gravity assembles structures in a bottom-up fashion. The haloes of dark matter acquire angular momentum via tidal torques (Peebles, 1969; Fall & Efstathiou, 1980) from interacting structures, and it is in these haloes, and their substructure, that all galaxies form. In Sect. 1.5 we discussed the virialization and structure of dark matter haloes. In the classical picture of galaxy formation, baryons are accreted spherically into the forming dark matter haloes. Being a collisional fluid, the baryons can not redistribute its kinetic energy into random motions, bur rather as *thermal energy* via accretion shocks. For the gas to be supported in the potential well the thermal energy must equal the gravitational potential energy ($U_{\rm therm} = U_{\rm grav}$), i.e. the gas must have heated to the *virial temperature*

$$T_{\rm vir} = \frac{2}{3}\frac{\mu m_{\rm H}}{k_{\rm B}}\frac{GM_{\rm vir}}{r_{\rm vir}}, \tag{1.86}$$

where μ is the mean molecular weight, $k_{\rm B}$ is the Boltzmann constant and $m_{\rm H}$ is the mass of a hydrogen atom. The Milky Way halo virial mass is today believed to be $M_{\rm vir} \sim 1 - 2 \times 10^{12}\,{\rm M}_\odot$ (Dehnen & Binney, 1998; Xue et al., 2008; Li & White, 2009), indicating that our galaxy is surrounded by a hot tenuous corona at $T \sim 10^6$ K. This leftover component from galaxy formation has been observed by e.g. Wang et al. (2005) as X-ray absorption lines of highly ionized species such as OVIII in the spectra of several active galactic nuclei. However, the extent and total mass of the hot gas halo is still unknown.

The Formation of Disks — In the naive scenario that all infalling gas is heated to the virial temperature, a galactic disk forms as gas cools down and condenses into the central parts of the dark matter halo, leading to a gravitationally unstable disk that fragments into dense star forming clouds. A realistic angular momentum content can be accounted for if most of the angular momentum is retained in the assembly process. In this picture, the host halo is responsible for the final galaxy characteristics. A simple model describing the properties of disk galaxies formed in a hierarchical Universe was developed by Mo et al. (1998). In this study, a rotationally supported disc with exponential surface density profile is assumed to form with a mass and angular momentum which are fixed fractions of those of its surrounding dark halo.

We assume that the final disks are thin and have exponential surface density profiles given by

$$\Sigma(r) = \Sigma_0 \exp(-r/r_{\rm d}), \tag{1.87}$$

where $r_{\rm d}$ and Σ_0 are the disk scalelength and central surface density, which in turn defines the disk mass

$$M_{\rm d} = 2\pi\Sigma_0 r_{\rm d}^2. \tag{1.88}$$

For simplicity we assume that the host dark halo is an isothermal sphere with a constant circular velocity $v_{\rm c}$ defined via the density profile (see also Eq. 1.69)

$$\rho(r) = \frac{1}{4\pi r^2}\frac{dM(r)}{dr} = \frac{v_{\rm c}^2}{4\pi G r^2}. \tag{1.89}$$

The angular momentum of the disk is

$$J_{\rm d} = 2\pi \int v_{\rm c}\Sigma(r) r^2 {\rm d}r = 2 M_{\rm d} R_{\rm d} v_{\rm c}, \tag{1.90}$$

and the angular momentum J of a dark matter halo can be defined via the dimensionless spin parameter λ (Peebles, 1969) as

$$\lambda = \frac{J|E|^{1/2}}{GM^{5/2}}, \tag{1.91}$$

where E is the total energy of the halo. We assume that the disk mass is a fixed fraction $m_{\rm d}$ of the halo mass, and similarly for the disk angular momentum ($J_{\rm d} = j_{\rm d} J$). The total energy of the (truncated) isothermal halo is, assuming circular orbits and using the virial theorem, $E = -Mv_\mathrm{c}^2/2$. Using this, and the relations above, we can express the scale radius as

$$r_\mathrm{d} = \frac{1}{\sqrt{2}} \left(\frac{j_\mathrm{d}}{m_\mathrm{d}} \right) \lambda r_{200}, \tag{1.92}$$

where we have used the fact that a halo of circular velocity v_c at redshift z have a virial radius and mass

$$r_{200} = \frac{v_\mathrm{c}}{10H(z)} \quad \text{and} \quad M = \frac{v_\mathrm{c}^2 r_{200}}{G} = \frac{v_\mathrm{c}^3}{10GH(z)}, \tag{1.93}$$

where $H(z)$ is given by Eq. 1.24. We can now express the disk characteristics r_d and Σ_0 in the following convenient way

$$r_\mathrm{d} \approx 8.8\, h^{-1}\,\mathrm{kpc} \left(\frac{\lambda}{0.05} \right) \left(\frac{v_\mathrm{c}}{250\,\mathrm{km\,s^{-1}}} \right) \left(\frac{H}{H_0} \right)^{-1} \left(\frac{j_\mathrm{d}}{m_\mathrm{d}} \right) \tag{1.94}$$

$$\Sigma_0 \approx 380\, h\,\mathrm{M_\odot}\,\mathrm{pc}^{-2} \left(\frac{m_\mathrm{d}}{0.05} \right) \left(\frac{\lambda}{0.05} \right)^{-2} \left(\frac{v_\mathrm{c}}{250\,\mathrm{km\,s^{-1}}} \right) \left(\frac{H}{H_0} \right) \left(\frac{m_\mathrm{d}}{j_\mathrm{d}} \right)^2, \tag{1.95}$$

where similar expression can be derived for arbitrary halo profiles. The equations above contain many unknowns, but good agreement with observed galactic scale radii (Courteau, 1997) and the luminosity-rotational velocity relation (Tully-Fisher relationship, Tully & Fisher, 1977) can be found (Mo et al., 1998) provided that

- The masses of discs are a few per cent of those of their haloes ($m_\mathrm{d} \leq 0.05$).
- The specific angular momenta of discs are similar to those of their haloes ($j_\mathrm{d} \approx 1$).
- Present-day disks were assembled recently (at $z \sim 1$).

As we will discuss in Chap. 6, these criteria turn out to be very difficult to fulfill in fully hydrodynamical simulations of galaxy formation, where a substantial fraction of the angular momenta is lost during disk assembly, leading to disks that feature scale radii that are smaller by a factor of a few in comparison to Eq. 1.94.

Modes of Accretion — For the gas to reach the halo virial temperature upon infall, it must have been heated by an expanding virial shock. However, it is not clear that stable shocks can persist in all haloes or at all cosmic epochs as stable extended shocks can only exist when the pressure in the post-shock gas is sufficient to balances the gravitational attraction toward the halo centre. Birnboim & Dekel (2003) (see also Dekel & Birnboim, 2006) studied the existence of accretion shocks in a numerical setting and provided an approximate analytical framework for shock stability under radiative cooling. The pressure of an adiabatic gas is given by

$$P = (\gamma - 1)\rho e, \tag{1.96}$$

where e is the specific internal energy and γ is the adiabatic index defined as $\gamma \equiv (\partial \ln P / \partial \ln \rho)_\mathrm{ad}$. At hydrostatic equilibrium, an adiabatic gas is gravitationally unstable to compression when the adiabatic index $\gamma < 4/3$. If we allow for radiative energy loss at a rate per unit mass q, γ should be replaced by[13]

$$\gamma_\mathrm{eff} = \gamma - \frac{\rho}{\dot{\rho}} \frac{q}{e}. \tag{1.97}$$

From this equation it is clear that any modification to the usual $\gamma > 4/3$ criterion is due to the magnitudes of the involved compressional and cooling timescales, give by

$$t_\mathrm{comp} \equiv \frac{\rho}{\dot{\rho}} \quad \text{and} \quad t_\mathrm{cool} \equiv \frac{e}{q} \tag{1.98}$$

respectively. A perturbation analysis of a spherical shell in hydrostatic equilibrium (Birnboim & Dekel, 2003) shows stability if

$$\gamma_\mathrm{eff} > \frac{2\gamma}{\gamma + 2/3} \equiv \gamma_\mathrm{crit}. \tag{1.99}$$

---

[13] Consider $\gamma_\mathrm{eff} = \frac{\mathrm{d}\ln P/\mathrm{d}t}{\mathrm{d}\ln \rho/\mathrm{d}t}$, Eq. 1.96 and the fact that the time derivative of the internal energy becomes $\dot{e} = -P\dot{V} - q$.

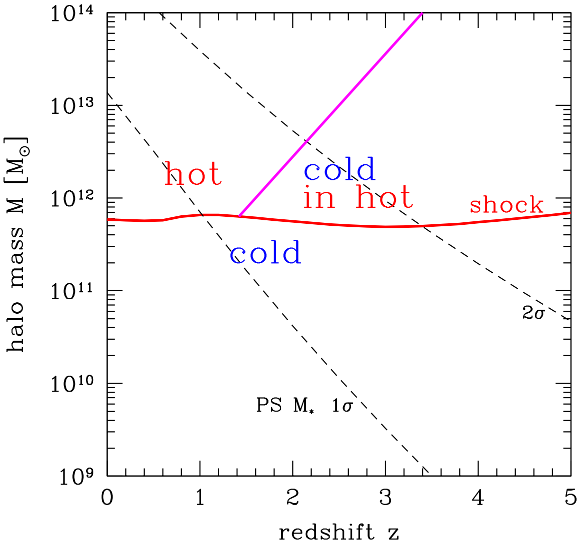

Figure 1.1: Cold streams and shock-heated medium as a function of halo mass and redshift. The nearly horizontal curve is the typical threshold mass for a stable shock in the spherical infall, below which the flows are predominantly cold and above which a shock-heated medium is present. The inclined solid curve is the upper limit for cold streams. The hot medium in massive haloes at high redshifts hosts cold streams which allow disk growth and star formation, while haloes of a similar mass at lower redshifts are all hot. Figure from Dekel & Birnboim (2006).

For a monoatomic gas, relevant for ionized gaseous haloes, $\gamma = 5/3$ which gives us the new stability criterion $\gamma_{\rm crit} > 10/7$. At a shocks, all Euler variables change discontinuously in a way given by the standard *Rankine-Hugoniot* shock jump conditions (see e.g. Shu, 1992; Toro, 1999). By considering the situation of a stalling shock, i.e. the case at which the post-shock gas pressure is able to precisely balance the gravitational pull, Eq.1.99 for a monoatomic gas, in conjunction with the jump conditions, can be be phrased as

$$\frac{\rho_0 r_{\rm s} \Lambda(T_1)}{|u_0|^3} < 0.0126 \tag{1.100}$$

where ρ_0 and u_0 is the pre-shock halo gas density and infall velocity, and T_1 the post-shock temperature, which here can be expressed as $T_1 = 3\mu u_0^2/16 k_{\rm B} N$. This simple relation allows us to evaluate shock stability given a cooling function $\Lambda(T)$ (e.g. Sutherland & Dopita, 1993) at a halo radius $r_{\rm s}$. Using the cooling function ($\Lambda = n^2 q$), and the fact that the internal energy $E = \frac{3}{2} n k_{\rm B} T$, we can phrase the cooling time in the more common form

$$t_{\rm cool} \equiv \frac{3 k_{\rm B} T}{2 n \Lambda(T)}, \tag{1.101}$$

where n is the particle number density. By using a halo model, e.g. the Press-Schechter formalism (c.f. Sect. 1.7) and the assumption of NFW mass profiles, Dekel & Birnboim (2006) carefully analyzed the cold and hot accretion regimes of galaxy formation as a function of dark matter halo mass, redshift and gas metallicity. In Fig. 1.1 we reproduce one of their main results: below a characteristic halo mass $M_{\rm vir} \sim 10^{11.7}\,{\rm M}_\odot$, given $Z = 0.1 Z_\odot$ and $r = 0.1\,R_{\rm vir}$, all gas is accreted as cold gas a all cosmic epochs. At high redshift ($z \gtrsim 2$), cold gas can co-exist with shock heated gas in the massive halos as *cold streams* of gas, while this is not expected to be the case at lower redshifts. Filamentary accretion is expected to occur as cosmic gas cools down into the potential well of the dark matter filaments constituting the cosmic web. This gas is later accreted onto the large galaxies situated in the nodes of the web, as well as onto smaller galaxies the exist within the filaments. This simple picture has been confirmed in the literature (e.g. Kereš et al., 2005, 2009; Brooks et al., 2009; Agertz et al., 2009), and in Chap. 5 we demonstrate how a phase of cold stream accretion in massive galaxies can explain modern observations of high redshift galaxies. As it turns out, Milky Way sized galaxies at $z \gtrsim 2-3$, the epoch at which star formation activity in the Universe peaks, acquire their baryons *mainly* via cold streams. This might be in stark contrast to the scenario of how angular momentum is acquired provided by the classical picture of disk formation outline in the previous section. The full implication of the bi-modal nature of gas accretion, and hence galaxy formation, is at the time of writing still being developed, but will undoubtedly remain an intriguing component of modern galaxy formation theory.

2.

Numerical Methods for Structure Formation

> 'But what is it good for?'
>
> — *Engineer at the Advanced Computing Systems Division of IBM, commenting on the microchip, 1968*

The previous chapter introduced the analytical framework of galaxy formation. To study cosmic structures past what we can be learned via perturbation theory, the full non-linear evolution must be solved numerically. This chapter outlines techniques commonly adopted for studying the evolution of collisionless $N-$body systems as well as self-gravitating fluids.

2.1. The Governing Equations

The state of a system of particles can be represented by the single-particle distribution function $f(\boldsymbol{x}, \boldsymbol{p}, t)$, where $\boldsymbol{x}$ and $\boldsymbol{p}$ are position and momentum respectively. At a time t, $f(\boldsymbol{x}, \boldsymbol{p}, t)\mathrm{d}^3x\mathrm{d}^3p$ equals the number of particles within the six-dimensional phase space volume $\mathrm{d}^3x\mathrm{d}^3p$ centered at $\boldsymbol{x}, \boldsymbol{p}$. By considering the flow in phase-space, it can be shown that $f(\boldsymbol{x}, \boldsymbol{p}, t)$ obeys the *Boltzmann equation*

$$\frac{\partial f}{\partial t} + \frac{\boldsymbol{p}}{m} \cdot \frac{\partial f}{\partial \boldsymbol{x}} - m\boldsymbol{\nabla}\phi \cdot \frac{\partial f}{\partial \boldsymbol{p}} = \left.\frac{\partial f}{\partial t}\right|_c \tag{2.1}$$

where the right hand side accounts for collisions, leading to non-conservative effects in the phase-space volume. In a self-gravitating system, the gravitational potential obeys *Poissons's equation*

$$\boldsymbol{\nabla}^2\phi = 4\pi G \int \mathrm{d}^3p f(\boldsymbol{x}, \boldsymbol{p}, t) = 4\pi G\rho. \tag{2.2}$$

Neglecting collisions ($\partial f/\partial t|_c = 0$), this equation is called the collisionless Boltzmann equation, or the *Vlasov equation*, and implies that phase space density is conserved along particle trajectories. If the initial phase space density of a system is completely known, then its evolution can, in principle, be found from the Vlasov equation. For collisionless systems at equilibrium with a very large number of particles, e.g. elliptical galaxies, it is possible to analytically describe the phase space density and solve the Vlasov equation. However, for the non-linear development of cosmic structure, this approach is not suitable; the number of dark matter particles is in practice infinite and the phase space structure becomes increasingly complex. The standard practice is to discretize phase-space using a set of N particles and evolve them according to the equations of motions

$$\frac{\mathrm{d}\boldsymbol{x}_p}{\mathrm{d}t} = \boldsymbol{v}_p \quad \text{and} \quad \frac{\mathrm{d}\boldsymbol{v}_p}{\mathrm{d}t} = -\boldsymbol{\nabla}\phi, \tag{2.3}$$

which if collisions are avoided satisfies the Vlasov equation and conserves phase space along the particle trajectories. Collisions are usually avoided by softening the gravitational interaction between particles, see Sect. 2.2.

The above discussion is valid for collisionless systems, e.g. dark matter where the timescale of collision is greater than the age of the Universe. This is of course not the case of gas, which is why a full hydrodynamical approach must be undertaken instead of a simple $N-$body approximation. By integrating over velocity moments of Eq. 2.1, we retrieve conservation laws for mass, momentum and energy: the equations of hydrodynamics. The zeroth order solution is called the *Euler equations*. These are the equations we solve in this thesis, which in vector notation are given by

(2.4) *Continuity:* $$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) = 0$$

(2.5) *Momentum:* $$\frac{\partial}{\partial t}(\rho \boldsymbol{v}) + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v} \otimes \boldsymbol{v} + P) = -\rho \boldsymbol{\nabla} \phi$$

(2.6) *Energy:* $$\frac{\partial}{\partial t}(\rho e) + \boldsymbol{\nabla} \cdot [\rho \boldsymbol{v}(e + P/\rho)] = -\rho \boldsymbol{v} \cdot \boldsymbol{\nabla} \phi,$$

where $e = E/\rho = (v^2/2 + U)$ is the specific total energy, U is the internal energy and $\otimes$ is the outer vector product[14]. The equations are closed by assuming an polytropic equation of state for the pressure, $P = (\gamma - 1)\rho U$, where γ is the adiabatic index ($\gamma = 5/3$ and $7/5$ for a mono and diatomic gas respectively). The gas temperature is provided by the ideal gas law $P = \rho k_{\rm B} T/\mu m_{\rm H}$, where μ is the mean molecular weight, $k_{\rm B}$ is the Boltzmann constant and $m_{\rm H}$ is the mass of a hydrogen atom. Higher order derivations of Eq. 2.4-2.6, via so called *Chapman-Enskog* procedure (c.f. Shu, 1992), leads to additional diffusive terms e.g. viscous stresses in the momentum equation and thermal conduction in the energy equation[15]. These effects are omitted in this thesis.

In reality, radiative cooling dissipates the gas internal energy. In the limit of an optically thin medium, the gas loses energy without depositing it to its surrounding. This process is modeled by augmenting Eq. 2.6 by a cooling function $\Lambda(\rho, T)$. Analogous to this, heating via e.g. a cosmic UV background field (e.g. Haardt & Madau, 1996) can also be modeled by a adding a source term $\Gamma(\rho, T)$.

2.2. N-body Methods

In the previous chapter we argued that the collisionless dark matter component dominates the matter content of our Universe, and that dark matter haloes are the nurseries for galaxy formation. On large scales, gravity dominates the evolution of structure, and baryonic effects only become important in the heart of dark matter haloes; in large galaxies, the baryons can contribute as much too the gravitational potential as the dark matter. Pure $N-$body simulations, i.e. neglecting the more complicated gas component, are therefor valuable tools for studying the non-linear evolution of cosmic matter. For example, the mass function of dark matter haloes and their substructure are believed to correlate with the luminosity function galaxies we see around us, hence providing strong constraints on a tested cosmological model.

At the time of writing, the largest simulations of large scale structure, i.e. on scales of several Gpc, use up to ~ 100 billion particles (e.g. Teyssier et al., 2009). Individual dark matter haloes have successfully been studied using on the order of billions of particles. Fig. 2.1 shows a Milky Way sized ($M_{200} \sim 10^{12}$ M$_\odot$) dark matter halo at $z = 0$. This simulation was carried out using the PKDGRAV tree code (Stadel, 2001), and reaches $\sim 1.3 \times 10^9$ particles within the virial radius.

Solvers — A multitude of solvers for the equations of motions of purely gravitational systems have been developed in the last four decades, not just in the field of astrophysics, but also plasma physics where charge is assigned to the particles instead of mass (c.f. Hockney & Eastwood, 1981). Reviewing these is beyond the scope of this thesis, and we summarize only the most common ones in the field of astrophysics.

The most straightforward way to integrate an $N-$body system is via direct summation; the so called Particle-Particle (PP) method. The force from particle j acting on particle i is

(2.7) $$\boldsymbol{F}_{ij} = \frac{Gm^2(\boldsymbol{x}_j - \boldsymbol{x}_i)}{(\epsilon^2 + |\boldsymbol{x}_i - \boldsymbol{x}_j|^2)^{3/2}},$$

where the particles are at positions $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ and they all have the same mass m. The parameter ϵ is called the *softening length*, and is necessary to avoid two-body scattering. While formally a correct approach, direct summation quickly becomes expensive when many particles are involved. For a system of $N_{\rm p}$ particles,

---

[14] Also referred to as a *Dyadic* product.

[15] These more general equations are called the *Navier-Stokes* equations.

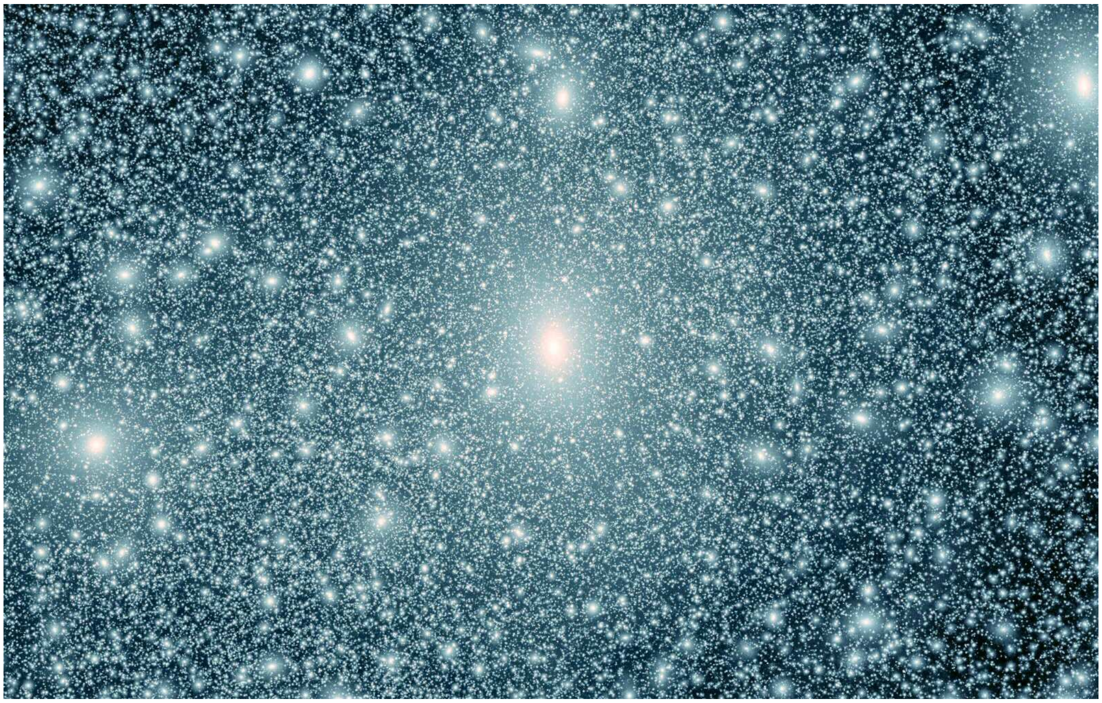

Figure 2.1: Density map of the inner 200 kpc of the billion particle GHALO Milky Way halo simulation (Stadel et al., 2009). Note the richness of substructure predicted within the ΛCDM cosmology. The abundance of the objects greatly surpasses the number of satellite galaxies detected so far within our own halo, indicating that galaxy formation becomes extremely inefficient on small scales. This *missing satellite problem* was pointed our by Moore et al. (1999) and Klypin et al. (1999).

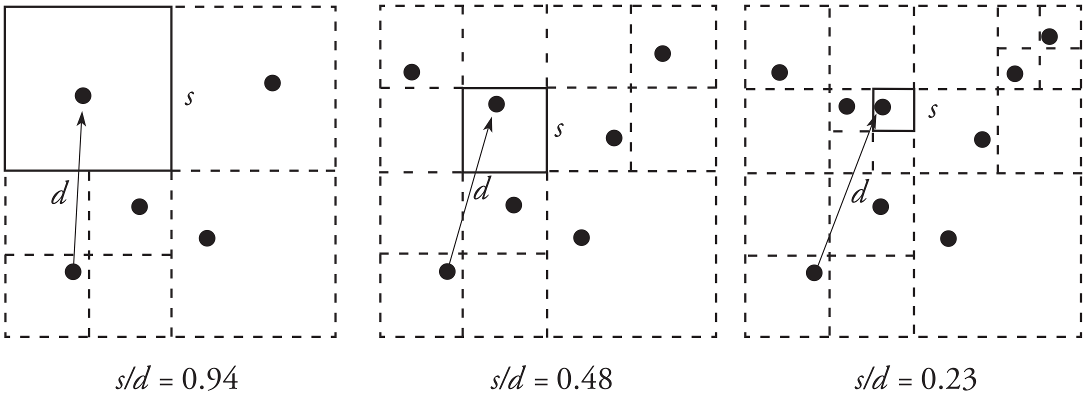

Figure 2.2: An illustration of how the opening angle $\theta = s/d$ set the tree accuracy. As θ is decreased from left to right, the tree cells are opened up, allowing for more particles to participate in the gravity calculation (from Pfalzner & Gibbon, 1996).

Eq. 2.7 must be evaluated $N_\mathrm{p}(N_\mathrm{p}-1)/2$ times, making the computational cost of the method to be of order $\mathcal{O}(N_\mathrm{p}^2)$. It is therefore only an effective method for small N_p or when the close-range dynamics is of great interest, e.g. star clusters. This numerical obstacle can be circumvented, and we here present the two most common ways of doing so.

☐ *Tree Codes* — The basic philosophy of a tree code is to calculate forces from close range particles using direct summation while far away particles are "smeared out" and treated as large pseudo-particles. For details of the method, c.f. Pfalzner & Gibbon (1996). The tree method was introduced by Barnes & Hut (1986), and this early implementation works as follows. The simulation box is subdivided into 2^D cells, where D is the dimensionality of the problem. If a cell contains more than one particle, it is subdivided. In this way, a tree structure (here an *oct-tree*) is constructed where the box is the root, each new sub-cell is a twig, and the final cells, containing only one or no particle, are leafs. The tree can hence be regarded as a book-keeping device for the particles. Once the tree is built, the center of mass is calculated for all tree cells. For each particle, the force calculation starts by traversing the tree, starting at the root. The size of a cell, s, is compared with the distance to the particle, d. The relation

$$\frac{s}{d} \leq \theta, \tag{2.8}$$

where θ is called the tolerance parameter (or opening angle), tells the code if the current level is detailed enough. If Eq. 2.8 is unfulfilled, the node is opened up and the tree structure is traversed further. This process is illustrated in Fig. 2.2. Once satisfied, the force is found using the previously calculated mass of the current twig or leaf. In this way, particles feel long distance forces only from larger pseudo-particles. A small θ makes the tree code act more like a PP code (time consuming but high force accuracy), while a large θ speeds up the code but decreases the force accuracy. The suitable value of θ depends on the problem at hand and must be calibrated to direct summation tests. When the force is found on all of the particles, they are moved according to the equations of motion and integrated in time.

Modern implementations of tree codes (e.g. Stadel, 2001; Springel, 2005) achieve higher accuracy compared to the original Barnes & Hut (1986) scheme by not only accounting for the center of mass of each tree cell, but higher order multipoles. The actual choice of tree structure is also arbitrary, and varies in the literature. The computational cost of the tree method depends on the choice of θ, but normally it is of the order of $\mathcal{O}(N_\mathrm{p} \log N_\mathrm{p})$, although it can be made as efficient as order $\mathcal{O}(N_\mathrm{p})$[16].

☐ *Mesh Based Codes* — The most standard way of using an Eulerian mesh for $N-$body methods is called the *Particle-Mesh* (PM) method. An outline of the method is as follows:

1 – Particle masses are assigned to the computational grid using an assignment function.
2 – Poisson's equation is solved on the grid in order to get the potential.
3 – The force field is calculated from the potential.
4 – The force is interpolated on the grid to find the force on each particle.
5 – The equations of motions are integrated over a chosen time step.

The first step is usually done in a way that reduces the force fluctuations when particles are close to each other, i.e. one demands "smoothness". This means that there will be a level of continuity in the derivatives of the approximated function used in assigning the masses to the grid. Let us adopt a notation where $\boldsymbol{x}_i$ is the position of the *i*th particle, $G = 1$, the length of the simulation cube and the mass is unity, N_g is the number of grid cells and the vector $\boldsymbol{q} = \boldsymbol{n}/N_\mathrm{g}^{1/3}$, where $\boldsymbol{n}$ is the grid position. The density on the grid is first calculated from

$$\rho(\boldsymbol{q}) = \frac{N_\mathrm{g}}{N_\mathrm{p}} \sum_{i=1}^{N_\mathrm{p}} W(\boldsymbol{x}_i - \boldsymbol{q}) \tag{2.9}$$

where W is the function that describes the way of assigning the mass to the grid. The most common ones in the literature are:

Nearest Grid Point (NGP): The density arising from a particle is assigned to the nearest grid point. This method is rather crude and is rarely used due to discontinuous force values.

Cloud In Cell (CIC): This is a better approximation than NGP, and involves a linear interpolation scheme to the 2^D grid points defining the cubical grid cell containing the particle, where D is the dimension of the problem, see Fig. 2.3. The CIC technique gives continuous force values but discontinuous first derivatives.

---

[16] This is achieved by using the Fast Multipole Method (FMM), an algorithm that reduces the computational cost of tree multipole calculations.

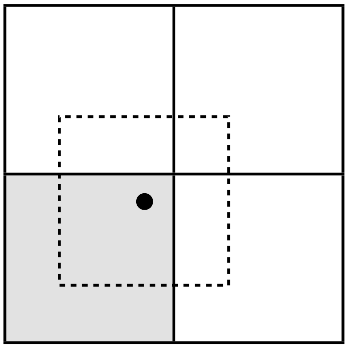

Figure 2.3: A 2D illustration of the Nearest-Grid-Point (NGP) and Cloud-In-Cell (CIC) techniques. In NGP, a particle (black dot) assigns all of its mass to the nearest cell (shaded area). In CIC, the mass is distributed to surrounding 2^D cells according to the area of the dotted square.

Triangular Shaped Cloud (TSC): This method uses a quadratic interpolation scheme. For $D = 3$, the particles will assign mass to the 27 nearest grid points.

In the second step, the PM method solves Poisson's equation (Eq. 2.3) by performing the sum

$$\phi(\boldsymbol{q}) = \frac{1}{N_{\rm g}} \sum_{\boldsymbol{q'}} \mathcal{G}(\boldsymbol{q} - \boldsymbol{q'})\rho(\boldsymbol{q'}), \tag{2.10}$$

where $\mathcal{G}$ is a Green's function for the Poisson equation given by

$$\mathcal{G} = -\frac{1}{4\pi|\boldsymbol{q} - \boldsymbol{q'}|}. \tag{2.11}$$

In the case of cosmological simulations where periodic boundary conditions are used, $\mathcal{G}$ is more complicated. The advantage of the PM method enters when calculating Eq. 2.10. This step is calculated using a Fast Fourier Transform (FFT) technique, which turns the convolution into a product of the Fourier transformed Green's function[17] and the transformed density field,

$$\hat{\phi}(\boldsymbol{k}) = \hat{\mathcal{G}}(\boldsymbol{k})\hat{\delta}(\boldsymbol{k}). \tag{2.12}$$

The potential is then transformed back to real space in order to obtain the cell-centered forces as

$$\boldsymbol{F}(\boldsymbol{q}) = -\frac{1}{N_{\rm p}} \boldsymbol{D}\,\phi, \tag{2.13}$$

where $\boldsymbol{D}$ represents a chosen finite differencing scheme. The next step is to interpolate the force back to the particle positions as

$$\boldsymbol{F}(\boldsymbol{x}_i) = -\sum_{\boldsymbol{q}} W(\boldsymbol{x}_i - \boldsymbol{q})\boldsymbol{F}(\boldsymbol{q}). \tag{2.14}$$

The assignment function W should here be the same as the density assignment scheme in step one for consistency. The main advantage of the PM scheme is speed. Because of the use of FFTs, the PM method shows a computational cost of the order $\mathcal{O}(N_{\rm p} + N_{\rm g} \log N_{\rm g})$. This is a substantial improvement for large $N_{\rm p}$ when compared to the PP method.

Currently, the "vanilla" PM method is only used in simulations of large scale structure where resolution on small scale is of less importance. Tree codes (e.g. Springel, 2005) often employ a PM method for integrating the evolution of the early (smooth) Universe. Modern mesh based codes employs adaptive or nested meshes to achieve higher resolution in regions of many particles. Here, multi-grid solvers or relaxations techniques must be employed to find the correct potential on the embedded multi-resolution grids.

---

[17] For a standard seven-point approximation to the Laplacian, $\hat{\mathcal{G}} \sim \left[\sin^2\left(\frac{k_x}{2}\right) + \sin^2\left(\frac{k_y}{2}\right) + \sin^2\left(\frac{k_z}{2}\right)\right]$

Time Integration — Most modern $N-$body codes evolve particles in time via so called *symplectic integrators*. Symplectic integration schemes can be derived from Hamilton's equations, which assures excellent conservation properties as phase-space volume is automatically conserved. The scheme is based on "kick" and "drift" steps, in which a "kick" refers to a momentum update keeping the position fixed, and a "drift" refers to a position update at a fixed momentum. By alternating kick and drift steps in increasingly elaborate sequences, higher-order integrators can be constructed. A commonly used scheme is the second order *Leapfrog* integrator in which the particles drift for $\Delta t/2$, kick for Δt and then drift again for $\Delta t/2$. This is called a *Drift-Kick-Drift* scheme and can formally be summarized as

$$\begin{aligned} \textit{Drift:} \quad x_{n+\frac{1}{2}} &= x_n + v_n \frac{\Delta t}{2} \\ \textit{Kick:} \quad v_{n+1} &= v_n + \nabla\phi_{n+\frac{1}{2}}\Delta t \\ \textit{Drift:} \quad x_{n+1} &= x_{n+\frac{1}{2}} + v_{n+1}\frac{\Delta t}{2}, \end{aligned} \tag{2.15}$$

where n is the integration step. An equally performing scheme is *Kick-Drift-Kick* Leapfrog, which is similar to Eq. 2.15 if but with the drift step acting on Δt and two kick steps on $\Delta t/2$.

Fully symplectic schemes are desirable but can not be achieved in methods involving variable time-stepping. This is the case for fully coupled $N-$body and hydrodynamic codes such as RAMSES, where each refinement level must allow for individual time-stepping, otherwise leading to instabilities in the hydro solver.

2.3. Numerical Hydrodynamics and Galaxy Formation

The development of modern computational fluid dynamics began with the advent of the first digital computer in the mid 20th century. Historically, numerical solutions of the fluid equation have involved finite element methods (FEM), or today more commonly finite difference or volume methods (FDM and FVM). These methods are general ways of solving partial differential equations, and the fluid equations in particular. Summarizing the historical background of numerical hydrodynamics is beyond the scope of this thesis, and we refer to standard textbooks (c.f. Chung, 2002).

While $N-$body methods have provided us with great insight to how cosmological structure formation proceeds, the study of actual galaxy formation must ultimately include the gaseous component to allow for star formation, enrichment of heavy elements etc. Due to the computational demand, simulations of galaxy formation have only included gas dynamical effects for roughly two decades. However, a hydrodynamical treatment of structure formation does not ensure the formation of realistic galaxies. Early numerical work by Navarro & Benz (1991), Katz (1992) and Navarro & White (1994) of galaxy formation in a cosmological context demonstrated the difficulty in forming objects that resembled observed galaxies. The galaxies turned out to be very compact and massive, and did not feature extended rotating disks, signaling a severe loss of angular momentum during the assembly process, as well as gas overcooling. It was then, and still is, unclear if these issues stemmed from numerical issues or a naive treatment of physical process e.g. supernovae, star formation, radiative effects etc. In Chap. 6 we discuss this issue in great detail and demonstrate what steps that need to be taken to reproduce realistic galaxies.

The equations of fluid dynamics were given in Eq. 2.4-2.6, and the way these equations are commonly solved, at least in the field of astrophysics, can in principle be split into two categories: *Lagrangian* and *Eulerian* schemes. Lagrangian methods solve the Lagrangian fluid equation (Eq. 1.31 and 1.32). This is often done by discretizing the fluid using particles, where fluid properties are acquired by smoothing over nearby neighbours. Fluid advection is achieved by moving the particles themselves according to the equation of motions in a completely mesh-free way. In contrast, Eulerian methods usually discretize the fluid equations on a mesh, on which fluid fluxes on adjacent cell boundaries are calculated to high precision, achieving fluid advection. Below, we outline the basic principles of both approaches.

Whilst both schemes aim to solve the same set of equation, the approximated discrete equations must differ and the simulation outcomes are not guaranteed to converge to the same results at all times. In Chap. 3 we analyze a set astrophysically relevant hydrodynamical problems and demonstrate this effect.

Smoothed Particle Hydrodynamics — Smoothed particle hydrodynamics (SPH) was first formulated as a tool for studying stellar structure (Gingold & Monaghan, 1977; Lucy, 1977). Today, SPH has firmly established itself as one the primary numerical tools in astrophysics, being actively used in the fields of planet formation (Mayer et al., 2002), star formation (Bate et al., 2003), galaxy formation (Katz et al., 1996; Governato et al., 2004) as well as engineering. Many versions of the method exists today (see reviews by e.g. Monaghan, 1992; Price, 2005), and we will here briefly outline the principles of its standard implementation.

The central idea of SPH is to represent a fluid by a set of discrete particles that move with the flow. In this Lagrangian frame, particle properties are updated by smoothing over their nearest neighbours, and fluid advection is achieved by simply moving the SPH particles. The basis of SPH is as follows. Consider the identity

$$A(\boldsymbol{r}) = \int A(\boldsymbol{r}')\delta(|\boldsymbol{r}-\boldsymbol{r}'|)d\boldsymbol{r}', \tag{2.16}$$

where A is an arbitrary variable and δ is the Dirac delta function. We can now approximate this relation by introducing a symmetric smoothing kernel W, with some characteristic width h, such that

$$\lim_{h\to 0} W(\boldsymbol{r}-\boldsymbol{r}', h) = \delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{2.17}$$

For consistency, the kernel W must also satisfy

$$\int W(\boldsymbol{r}-\boldsymbol{r}', h)\mathrm{d}\boldsymbol{r}' = 1. \tag{2.18}$$

By replacing δ with W in Eq. 2.16, this equation can be discretized onto a set of N particles by replacing the integral by a summation, i.e.

$$\begin{aligned} A(\boldsymbol{r}) &= \int \frac{A(\boldsymbol{r}')}{\rho(\boldsymbol{r}')} W(|\boldsymbol{r}-\boldsymbol{r}'|, h)\rho(\boldsymbol{r}')d\boldsymbol{r}' + O(h^2) \\ &\approx \sum_{b=1}^{N} m_b \frac{A_b}{\rho_b} W(|\boldsymbol{r}-\boldsymbol{r}_b|, h), \end{aligned} \tag{2.19}$$

where the mass element ρdV was replaced with the particle mass m, and A_b refers to the fluid quantity evaluated at the position of particle b. The summation interpolant in Eq. 2.19 shows the underlying philosophy of SPH and can be considered as the basis for all SPH methods. A strength of this formulation is that variable derivatives can be approximated as

$$\nabla A(\boldsymbol{r}) \approx \sum_b m_b \frac{A_b}{\rho_b} \nabla_a W_{ab}, \tag{2.20}$$

where $\nabla_a \equiv \partial/\partial \boldsymbol{r}_a$ and $W_{ab} = W(|\boldsymbol{r}_a - \boldsymbol{r}_b|, h)$. This means that derivatives can be evaluated via the analytically known kernel derivate. Note that we here assume $h \neq h(\boldsymbol{r})$ to simplify derivations. If varying smoothing lengths are allowed, as in modern usage of SPH, additional terms appear. For detailed derivations of relevant SPH operations (divergence, curl etc.) and their error properties, see Price (2005).

If we consider $A(\boldsymbol{r}) = \rho(\boldsymbol{r})$ in Eq. 2.19, we get the SPH expression for the density estimator

$$\rho_a = \sum_b m_b W_{ab} \tag{2.21}$$

The Lagrangian time derivative of the density estimator gives us the continuity equation

$$\frac{\mathrm{d}\rho_a}{\mathrm{d}t} = \sum_b m_b \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab}, \tag{2.22}$$

where we have used the relation $\mathrm{d}W_{ab}/\mathrm{d}t = \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab}$. Some flavours of SPH explicitly evolve the continuity equation, but if particles don't transfer mass to each other, Eq. 2.21 is a sufficient (integral form) of this equation.

The equations of motion can be derived from the fluid Lagrangian. This is a standard approach in the literature as the resulting equations assure excellent conservation properties. The fluid Lagrangian is, disregarding self-gravity, given by

$$\mathcal{L} = T - V = \int \left(\frac{1}{2}\rho v^2 - \rho u \right) \mathrm{d}V, \tag{2.23}$$

where T and V are the kinetic and internal energy respectively, and u is the specific internal energy. The SPH discretization of this equation is

$$\mathcal{L} = \sum_b m_b \left[\frac{1}{2}\boldsymbol{v}_b^2 - u_b(\rho_b, s_b) \right], \tag{2.24}$$

and the equations of motions are obtained by solving the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_a} \right) - \frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_a} = 0. \tag{2.25}$$

The derivatives become

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_a} = m_a \boldsymbol{v}_a \tag{2.26}$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_a} = \sum_b m_b \left.\frac{\partial u_b}{\partial \rho_b}\right|_s \frac{\partial \rho_b}{\partial \boldsymbol{r}_a} \tag{2.27}$$

The first law of thermodynamics[18] gives us $du = P d\rho/\rho^2$. Using this and the radial derivative of Eq. 2.21, Eq. 2.27 becomes

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_a} = -\sum_b m_b \frac{P_b}{\rho_b^2} \sum_c m_c \nabla_a W_{bc}(\delta_{ba} - \delta_{ca}) = -m_a \sum_b \left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} \right) \nabla_a W_{ab}. \tag{2.28}$$

Inserting this relation and Eq. 2.26 into the Euler-Lagrange equations gives us the SPH momentum equation

$$\frac{d\boldsymbol{v}_a}{dt} = -\sum_b \left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} \right) \nabla_a W_{ab}. \quad \textit{(Momentum equation)} \tag{2.29}$$

The energy equation is easily derived from the first law of thermodynamics as

$$\frac{du_a}{dt} = \frac{P_a}{\rho_a^2}\frac{d\rho_a}{dt} = \frac{P_a}{\rho_a^2} \sum_b m_b \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab}, \quad \textit{(Energy equation)} \tag{2.30}$$

where the last equality made use of the continuity equation (Eq. 2.22). Eq. 2.21, 2.29 and 2.30, together with an equation of state $P(\rho)$ for closure, summarizes the governing equations of the most basic SPH scheme.

More general forms of the discrete SPH equations can be obtained by directly discretizing the Euler equations. By doing so, the intrinsic errors of the method can be reduced by e.g. carefully choosing appropriate weighting functions (Read et al., 2010). In addition, dissipative terms is commonly added to the SPH equation to allow for more accurate treatment of shocks (artificial viscosity) and steep density gradients (artificial thermal conduction). A thorough discussion of this can be found in Price (2008).

Eulerian Hydrodynamics — We have on purpose written the Euler equations in Eqs. 2.4-2.6 as a series of conservation laws in the form

$$\frac{\partial \boldsymbol{U}}{\partial t} + \frac{\partial \boldsymbol{F}(\boldsymbol{U})}{\partial x} = \boldsymbol{S}, \tag{2.31}$$

where $\boldsymbol{U}$ is the state vector and $\boldsymbol{F}$ the flux vector given by, in one dimension for simplicity,

$$\boldsymbol{U} = \begin{pmatrix} \rho \\ \rho v \\ E \end{pmatrix} \quad \text{and} \quad \boldsymbol{F} = \begin{pmatrix} \rho v \\ \rho v^2 + P \\ v(E+P) \end{pmatrix}. \tag{2.32}$$

The right hand side of Eq. 2.31 includes source terms in the momentum and energy equation but we neglect those for the moment ($\boldsymbol{S} = 0$). In a finite volume scheme, the state of a system is discretized onto a mesh

---

[18] The First law states that no energy can be created or destroyed, just converted. Hence $dU = dQ - PdV$, where $dQ = 0$ for an adiabatic process. In specific quantities (per mass units), $V \rightarrow 1/\rho$, hence $dV \rightarrow (1/\rho) = -d\rho/\rho^2$

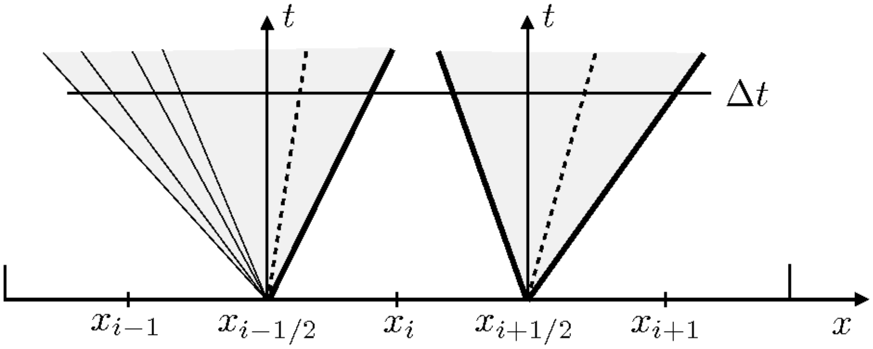

Figure 2.4: A one-dimensional illustration of a finite volume mesh discretization. State variables U are stored at cell centers x_i while fluxes are calculated at cell boundaries $x_{i\pm1/2}$. In Godunov schemes, fluxes are calculated by solving a Riemann problem for adjacent cells, e.g. $\mathcal{RP}[U_i, U_{i+1}]$. In this example illustration, the three different characteristic solutions to the Riemann solver are shown: shocks (thick lines), contact waves (dashed lines) and rarefaction waves (thin lines). The Riemann solution space, called the "star" region, is indicated in gray. No waves are allowed to overlap, and the time step Δt must hence be chosen to satisfy this.

with cell size Δx, see Fig. 2.4. The finite volume approximation of a continuous state variable at a time t_n in cell i is

$$U_i^n = \frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}} U(x, t^n)dx. \tag{2.33}$$

Using this in Eq. 2.31, it is possible to write the exact evolution of a volume averaged quantities as

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{F_{i+1/2}^{n+1/2} - F_{i+1/2}^{n-1/2}}{\Delta x} = 0, \tag{2.34}$$

where Δt the time step. The main problem is now how to calculate accurate fluxes at cell interfaces, and how to make the scheme stable and conservative.

The dynamics of the Euler equations is purely a matter of the propagation of signals, of which there are two kinds: sound waves and actual fluid advection. It is possible to rewrite Eq. 2.31 into a set of advection equations[19] by putting the pressure terms, being responsible for the sound waves, on the right hand side, hence treating them as source terms. In this way, one can solve a standard advection problem on the conserved scalar variables (ρ, ρv and E) and then add the source terms back at the end of the time step. Historically, this operator splitting method was the standard approach. Although a successful way of solving the Euler equations, or any kind of hyperbolic PDE, this is not formally correct. Formally, advection must be done in the eigen-frame of the system with the system's characteristics as advection velocities. This will assure a correct advection of not only the contact wave (i.e. the gas velocity), but also sound waves.

Using the chain rule, the conservations laws, here the Euler equations, can be written as an advection equation

$$\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{A}(\boldsymbol{U})\frac{\partial \boldsymbol{U}}{\partial x} = 0, \tag{2.35}$$

where the $\boldsymbol{A}(\boldsymbol{U}) = \partial \boldsymbol{F}/\partial \boldsymbol{U}$ is the *Jacobian matrix*. The eigenvalues of this matrix are the *characteristics* of the system and physically represent the speed of information propagation along the system's eigenvectors. By definition, $\boldsymbol{A} = \boldsymbol{K}\Lambda\boldsymbol{K}^{-1}$, where $\boldsymbol{K}$ is the matrix with columns of eigenvectors and Λ the diagonal matrix of eigenvalues. Eq. 2.35 can be decoupled into an independent set of variables $\boldsymbol{W} = \boldsymbol{K}^{-1}\boldsymbol{U}$, which give us the *characteristic form* of Eq. 2.31,

$$\frac{\partial \boldsymbol{W}}{\partial t} + \Lambda\frac{\partial \boldsymbol{W}}{\partial x} = 0. \tag{2.36}$$

The goal of an Eulerian scheme is identify the characteristic signal velocities in Λ, advect the state vector $\boldsymbol{W}$, and then reconstruct the original Euler variables $\boldsymbol{U}$ at the end of the time step. The most common way of doing this on a discrete mesh is via a *Godunov scheme* (Godunov, 1959). Here the characteristic wave analysis is achieved via *Riemann solvers* at each cell interface. A Riemann problem is an initial value

---

[19] For an advection velocity a, linear advection of a variable u is given by $\partial u/\partial t + a\partial u/\partial x = 0$.

problem where the governing equation of the form given in Eq.2.35, and the initial domain is separated into a left-hand ($\boldsymbol{U}_{\rm L}$) and a right-hand ($\boldsymbol{U}_{\rm R}$) state. In the discretized case above, a Riemann solver identifies the waves emerging at these characteristic wave speeds from the evaluation point $x_{i+1/2}$, and evaluates the boundary state $U^*_{i+1/2}(0)$. Depending on the physical input, waves can be either contact waves, shock waves or rarefaction waves, see Fig. 2.4. The intercell flux $\boldsymbol{F}_{i+1/2}$ is hence defined by

$$\boldsymbol{F}_{i+1/2} = \boldsymbol{F}(U^*_{i+1/2}(0)), \tag{2.37}$$

where $U^*_{i+1/2}(0)$ is the solution of the Riemann problem $\mathcal{RP}[U_i, U_{i+1}]$ evaluated at the cell boundary, i.e. $x/t = 0$ as it coincides with the t-axis in the local solution frame. There are no exact Riemann solutions in closed form for the 3D Euler equations, not even for the much simpler isentropic or isothermal equations. However, iterative schemes exist where arbitrary accuracy can be achieved. For a detailed treatment of the Riemann problem and it general solution, see Toro (1999).

The original Godunov scheme considered piecewise constant states within the cells. To provide higher resolution of discontinuities, Godunov's scheme can be extended to use piecewise linear approximations of the left and right-hand Riemann states in for each cell (van Leer, 1979), which results in a central difference scheme that is second-order accurate in space. Even higher order reconstruction e.g. the piecewise parabolic method (PPM, Colella & Woodward (1984)) exists but requires large stencil for the hydro calculation. In these higher order schemes, the extrapolated cell boundary values can sometimes overshoot close to steep gradients and shocks, causing oscillating solutions. In fact, Godunov (1959), proved that linear schemes for solving PDEs can only be at most first-order accurate to avoid the creation of new local minima or maxima. This property is called total variation diminishing (TVD). By introducing *flux limiters* to the discrete form of $F^{n+1/2}_{i+1/2}$, these oscillations can be removed, making the methods TVD.

The above discussion gives a flavour of how Eulerian methods are constructed. In a galaxy formation simulation, the method needs to include gravitational source terms coupled to the $N-$body component (stars and dark matter), adaptively refining meshes, gas cooling and sub-grid physical models. A complete description of the involved algorithms is beyond the scope of this thesis, and we refer to popular implementations for more information e.g. RAMSES (Teyssier, 2002) and ENZO (O'Shea et al., 2004).

2.4. Modeling Physical Processes

The previous sections outlined methods for solving $N-$body and hydrodynamical processes. While essential for studying the non-linear process of galaxy formation, these methods fail to account for physical processes such as the birth, life and death of stars, gas cooling, chemical evolution, radiative processes, feedback from young stars and supernovae etc. Gas cooling and heating can be accounted for in the energy equation (Eq. 2.6) and advection and mixing of different elements can be modeled as additional passive scalars in the Euler equations. However, this is not the case for the other examples given above. These are highly unresolved phenomena which require so called *sub-grid modeling*; simplified local treatments put in by hand, based on the information in, for mesh based codes, one or at most a few cells. Some examples of sub-grid models in simulations of galaxy formation are given below.

Star Formation — In a cosmological context, where simulation domains typically range from a few tens to hundreds of Mpc, state-of-the-art simulations (at the time of writing) are limited to cell sizes of $\Delta x \sim 10\,\mathrm{pc}$ in regions of maximum refinement (Agertz et al., 2009). This is the typical size of giant molecular clouds (GMCs). These structures consist of predominantly dense molecular gas with typical masses $\sim 10^5 - 10^6\ \mathrm{M}_\odot$ (Blitz et al., 2007), and are the locations in which most of the star formation takes place in galaxies. A self-consistent formation and destruction of GMCs is hence beyond the reach of current simulations.

Observationally, the star formation rate surface density in disks is often found to correlate well with the surface density of gas (Schmidt, 1959; Kennicutt, 1998). This correlation is usually parametrized as

$$\Sigma_{\rm SFR} \propto \Sigma^n_{\rm gas}, \tag{2.38}$$

where n has been found to range from $1 - 3$ depending on gas phase, size of observed region etc., see Chap. 6 for an extensive discussion on star formation in disk galaxies. In numerical simulations, the observed projected quantities are of little use as the local cell (or particle) quantity is a physical density. The most common way of modeling the conversion of gas into stars is to remove gas mass from a cell and convert it into a collisionless star particle according to a Schmidt-law of the form

$$\dot{\rho}_{\rm gas} = -\epsilon \frac{\rho_{\rm gas}}{t_{\rm dyn}} \text{ for } \rho > \rho_0, \tag{2.39}$$

where $\rho_{\rm gas}$ is the gas density, $t_{\rm dyn}$ is some dynamical time $\sim \rho_{\rm gas}^{-1/2}$, ϵ is the star formation efficiency and ρ_0 is the threshold for star formation. In later chapters we adopt the gas free-fall time, $t_{\rm dyn} = t_{\rm ff} = \sqrt{3\pi/32G\rho_{\rm gas}}$. Modern high-resolution simulations attempt to resolve the ISM well enough to motivate a high star formation threshold of $\gtrsim 10^2\,{\rm cm}^{-3}$, the typical average density of GMCs. Once a threshold is chosen, ϵ must be calibrated to render galaxies with realistic *projected* star formation relations (Eq. 2.38). As we demonstrate in Chap. 6, this simplified approach can be inappropriate for correctly reproducing realistic spiral galaxies. We note that there are alternative formulation of gas to star conversion laws in the literature (see e.g. Leroy et al. (2008) for a comprehensive summary). However, they are all designed to fit an observed relation, and hence makes use of a normalization constant similar to ϵ.

Supernovae Feedback — In stars with masses of $\sim 8 - 40\ {\rm M}_\odot$, fusion of heavy elements is sufficient to stabilize the stars at the end of their life cycles. Once the core starts forging the most tightly bound elements, such as nickel-56 (leading to iron-56 via radioactive decay), energy cannot be produced at the core via fusion. This leads to the formation of a massive iron core which later undergoes collapse, producing a shock wave that disintegrates the star and pollutes the interstellar gas with heavy elements. These cataclysmic events are called supernovae Type II events (SNII).

In stars with less than eight solar masses, the carbon produced via helium fusion does not fuse, and the star gradually cools down and becomes a white dwarf. White dwarf stars may then become Type Ia supernovae (SNIa) via mass accretion from a companion star, provided it ends it life as a red giant star. This forces the white dwarf mass over the *Chandrasekhar limit*[20] at $\sim 1.4\ {\rm M}_\odot$ leading to collapse.

The mass resolution in cosmological simulations of galaxy formation is orders of magnitude larger than actual stellar masses. This means that the formed star particles do not represent individual stars, but actual stellar populations with an associate initial stellar mass function (IMF). Given a mass function, it is possible to calculate the amount of stellar mass that ends its life as SNII or SNIa events at each time-step, see Chap. 6 for details. The energy associated with the SN blastwave is $\sim 10^{51}\,{\rm ergs}$ for a $10\ {\rm M}_\odot$ star (McKee & Ostriker, 1977a). A great deal of effort has been made to simulate this energy injection into the ISM (Katz et al., 1996; Stinson et al., 2006; Scannapieco et al., 2008). Many authors simple dump the total mount of energy released at each simulation time-step as thermal energy in a single cell, or SPH smoothing kernel, and lets the SN blast-wave develop self-consistently. Early simulation found that this led to a very rapid energy loss via radiative cooling in dense regions of star formation. This issue is often circumvented by shutting off cooling locally in fluid elements associated with the SN events (e.g. Gerritsen, 1997). Alternatively, the energy can be released directly as kinetic energy in fluid elements surround the SN region.

Models of supernovae explosion must be calibrated in conjunction with the adopted star formation model to ensure that a realistic star formation relation is reproduced. In addition, the amount of hot gas in galaxies like our Milky Way does not exceed a mass fraction of at most a few per cent (Ferrière, 2001), which constrains the models.

The Jeans Length — When gravitationally unstable regions are resolved by only a few cells, the true pressure gradients that could prevent gravitational collapse are not resolved, leading to *numerical fragmentation*. In addition, noise seeded at small scales can lead to the onset of instabilities, as the smallest scales have the shortest growth times (see Eq. 1.42). Numerical simulations by Truelove et al. (1997) demonstrated that the properties of gravitational collapse in isothermal disks converge if the Jeans length (Eq. 1.39) is resolved by more than four cells. This can be achieved by either refining the computational mesh based on this criteria,

---

[20] The Chandrasekhar mass is the mass at which a star no longer can balance self-gravity with degenerate electron-pressure (Shapiro & Teukolsky, 1983).

or artificially increasing the local sound speed in the affected cells. The latter solution can be implemented as a polytropic equation of state

$$T = T_0 \left(\frac{\rho}{\rho_0} \right)^{\gamma_0 - 1} . \tag{2.40}$$

Here, T_0 is set to be the cooling floor of the simulations which typically is ~ 300 K in simulations allowing for metal line cooling. At this temperature, the Jeans length reads

$$\lambda_{\mathrm{J}} \approx 300 \sqrt{1/n}\, \mathrm{pc}, \tag{2.41}$$

where n is the density it units of cm^{-3}. By requiring $4\Delta x = \lambda_{\mathrm{J}}$, where Δx is the finest grid resolution in a simulation, we can solve for the density, and hence ρ_0 in Eq. 2.40. By setting $\gamma_0 = 2.0$, the local temperature will be increased linearly with density, and the Truelove criteria will be satisfied.

Sketch of turbulent wakes by Leonardo da Vinci (1509).

3.

Fundamental Differences Between SPH and Grid Methods[21]

We have carried out a comparison study of hydrodynamical codes by investigating their performance in modeling interacting multiphase fluids. The two commonly used techniques of grid and smoothed particle hydrodynamics (SPH) show striking differences in their ability to model processes that are fundamentally important across many areas of astrophysics. Whilst Eulerian grid based methods are able to resolve and treat important dynamical instabilities, such as Kelvin-Helmholtz or Rayleigh-Taylor, these processes are poorly or not at all resolved by existing SPH techniques. We show that the reason for this is that SPH, at least in its standard implementation, introduces spurious pressure forces on particles in regions where there are steep density gradients. This results in a boundary gap of the size of an SPH smoothing kernel radius over which interactions are severely damped.

3.1. Introduction

The ability to numerically model interacting fluids is essential to many areas of astrophysics and other disciplines. From the formation of a star and its proto-planetary disk to galaxies moving through the intra-cluster medium, dynamical instabilities such as Kelvin-Helmholtz (KH) and Rayleigh-Taylor (RT) play a fundamental role in astrophysical structure formation. Most popular hydrodynamical methods can be divided into two classes: techniques following the gas using Eulerian grids (e.g. Laney, 1998; Leveque, 1998) and those which follow the Lagrangian motions of gas particles such as ‘Smoothed Particle Hydrodynamics’ (SPH) (Monaghan, 1992). Grid based techniques solve the fluid dynamical equations by calculating the flux of information through adjacent cell boundaries, while SPH techniques calculate the gas properties on each particle by averaging over its nearest neighbours. Due to the extensive use, and sometimes discrepant results of these techniques, it is interesting to carry out code comparison studies on well defined problems that test their ability to follow the basic gas physics they are designed to simulate. Recent code comparisons have been focusing on differences in a cosmological context (e.g. Frenk & et al, 1999; O’Shea et al., 2005; Regan et al., 2007). They all find differences between grid and SPH codes but due to the complexity of these types of simulations it is not obvious *how* the differences arise. Similarly, while SPH studies of galaxy-intracluster medium interactions by Abadi et al. (1999) found that only half the inter-stellar medium was removed from the galaxy. Using a grid based calculation with the same initial conditions, Quilis et al. (2000) found that all the gas could be removed and attributed the difference to the high resolution shock capturing ability of their Eulerian code. However we are not aware of a direct comparison between simulation methods in this context. Differences were found in the literature between different studies of the same problem.

Our test problem is to follow cold dense gas cloud moving through a low density hot medium. This is specifically designed to capture the same physical processes that occur during the formation and evolution of astrophysical structures. We will also study the shearing motion of two fluids of different densities to elucidate the problems that we find with this test. Similar configurations, including shock wave interaction

---

[21] This chapter has been published as Agertz et al. (2007) in Monthly Notices of the Royal Astronomical Society, Volume 380, Issue 3, pp. 963-978.

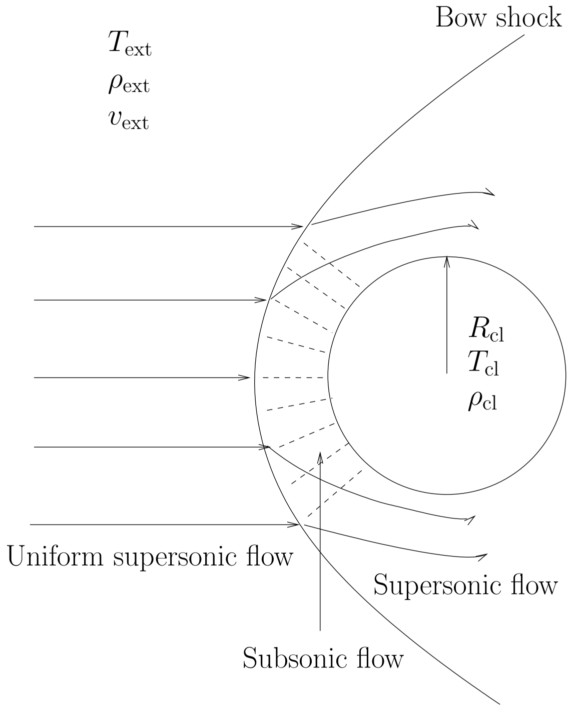

Figure 3.1: Illustration of the blob test. The external medium, which initially is in pressure equilibrium with the cloud, travels with a supersonic velocity creating a bow shock in front of the cloud. The post shock flow is subsonic until the smooth flow accelerates and again obtains supersonic speed on the lateral sides of the cloud.

with clouds, have been studied by e.g. Murray et al. (1993), Klein et al. (1994), Mac Low et al. (1994), Mac Low & Zahnle (1994), Vietri et al. (1997), Mori & Burkert (2000).

The paper is organized as follows. In Section 3.2 we briefly describe the main features of the test problem followed by analytical expectations in Section 3.3. In Section 3.4 we describe our numerical implementation of the setup as well as all codes used in our comparison. In Section 3.5 we present the results of our simulations followed by an explanation of the found discrepancies in Section 3.6. In Section 3.7 we summarize our results and briefly discuss their implications.

3.2. The Blob Test

A schematic view of the blob test problem can be seen in Fig. 3.1. A spherical cloud of gas is placed in a wind tunnel with periodic boundary conditions. The ambient medium is ten times hotter and ten times less dense than the cloud so that it is in pressure equilibrium with the latter. We will refer to this initial density contrast between the cloud and the medium as $\chi_{\rm ini}$. All of the gas is atomic hydrogen with molecular weight $\mu = 1.0$ and an adiabatic index $\gamma = 5/3$.

This setup is useful to investigate how different simulation codes handle typical astrophysical processes important for multi-phase systems, such as ram-pressure stripping and fragmentation through KH and RT instabilities.

3.3. Analytical Expectations

Although the nonlinear stages of the KH and RT instabilities cannot be fully described analytically, we can still use analytic arguments to estimate the characteristic disruption timescale for the cloud.

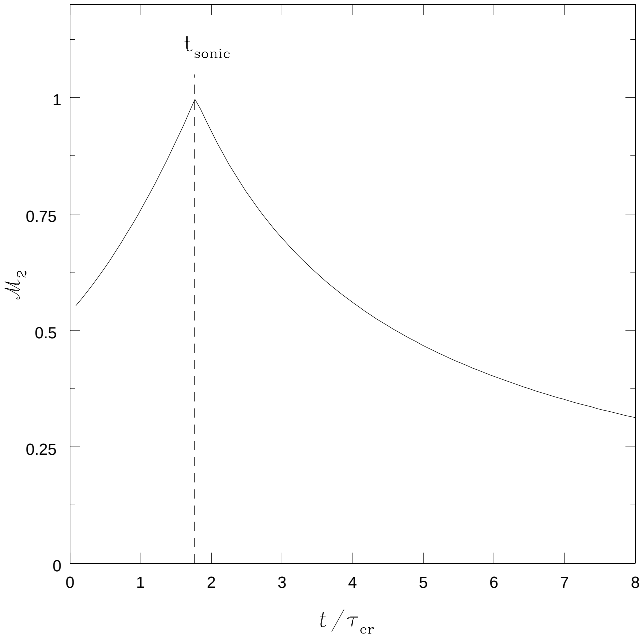

Figure 3.2: We plot the Mach number of the flow directly downstream of the shock on the symmetry axis of the cloud. The flow speed increases due to the weakened shock strength up to $t_{\rm sonic}$ where the relative motion of the cloud and wind turns subsonic.

In order to specify our problem we characterize the external medium with a sound speed $c_{\rm s}$ and assign it an initial velocity $v_{\rm ext} = \mathcal{M}c_{\rm s}$ with Mach number $\mathcal{M} = 2.7$. Furthermore, we place the cloud initially at rest in the computational domain. Since the wind is supersonic, a bow shock will form in front of the cloud with the post shock properties given by the Rankine-Hugoniot shock jump conditions. Because the cloud is accelerated by the wind, we will from now on perform all of our calculations *in the rest frame of the bow shock*, referring to pre-shock quantities with the subscript 1 and post-shock with 2. The shock conditions for the density, velocity and Mach number are (e.g. Shu, 1992)

$$\frac{\rho_2}{\rho_1} = \frac{v_1}{v_2} = \frac{(\gamma+1)\mathcal{M}_1^2}{(\gamma+1)+(\gamma-1)(\mathcal{M}_1^2-1)}, \tag{3.1}$$

$$\mathcal{M}_2^2 = \frac{2+(\gamma-1)\mathcal{M}_1^2}{2\gamma\mathcal{M}_1^2-(\gamma-1)}. \tag{3.2}$$

Formally we would take the obliqueness of the bow shock into account but for simplicity we will only consider the flow that enters at the symmetry axis of the cloud.

The cloud acceleration can be approximated by considering the maximum area that can gain momentum from the ambient flow. This implies that all gas in a cylinder in front of the cloud transfers momentum leading to an acceleration

$$a_{\rm cl} \sim \dot{v}_1 \sim \frac{\rho_{\rm ext}\pi R_{\rm cl}^2 v_1^2}{M_{\rm cl}}. \tag{3.3}$$

Integrating this equation leads us to the evolution of the pre-shock velocity

$$v_1(t) = \frac{l}{(t+l/v_{\rm ext})}, \tag{3.4}$$

where l is a characteristic length given by $l = M_{\rm cl}/2\pi R_{cl}^2\rho_{\rm ext}$. By using Eq. 3.4 to calculate the pre-shock Mach number together with Eq. 3.2 we can obtain a qualitative understanding of the post-shock velocity.

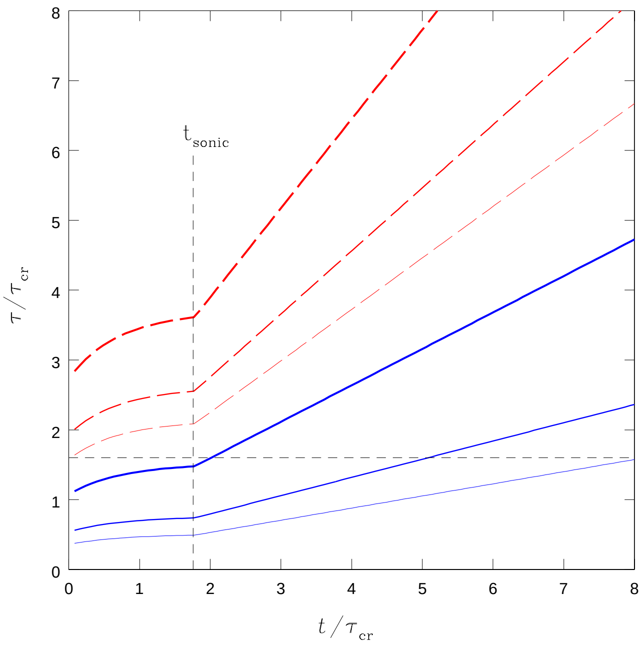

Figure 3.3: The time dependence of the growth rates of KH (solid, blue lines) and RT (dashed, red lines) instabilities. The lines represent different sizes of perturbation wavelengths: R_{cl} (thick), $R_{\mathrm{cl}}/2$ (middle) and $R_{\mathrm{cl}}/3$ (thin).

This velocity is crucial for the stability of the cloud surface and, as we will show in Section 3.3, for the destruction of the cloud itself. The evolution of the post-shock Mach number $\mathcal{M}_2$ is given by

$$\mathcal{M}_2^2 = \begin{cases} \frac{2+(\gamma-1)(v_1/c_s)^2}{2\gamma(v_1/c_s)^2-(\gamma-1)} & \text{for } t < t_{\mathrm{sonic}} \\ (v_1/c_s)^2 & \text{for } t > t_{\mathrm{sonic}} \end{cases} \tag{3.5}$$

Here t_{sonic} is the time at which $\mathcal{M}_1 = \mathcal{M}_2 = 1$ and the shock disappears. After this point, gas freely streams towards the cloud and the Mach number decreases only due to the continued acceleration. Notice that for $t < t_{\mathrm{sonic}}$, $\mathcal{M}_2 < 1$, even for $\mathcal{M}_1 = v_1/c_s \to \infty$. This means that behind the shock, the flow will always be subsonic and we expect instabilities to grow there. For $t \to \infty$, $\mathcal{M}_2 \to 0$ and the cloud will eventually be co-moving with the background flow. The evolution of the post-shock Mach number is shown in Fig. 3.2 in terms of the so-called "crushing time" defined as, in our notation,

$$\tau_{\mathrm{cr}} = \frac{2R_{\mathrm{cl}}\chi^{1/2}}{v_1}, \tag{3.6}$$

where χ is the density contrast between the cloud and the external medium. This is a natural timescale supersonic cloud evolution. We will naively use $\chi = \chi_{\mathrm{ini}} = 10$ and $v_1 = v_{\mathrm{ext}}$, representing our initial condition. During the interval of τ_{cr} a bow shock is formed and the shocked gas will form a smooth flow around the cloud, reaching supersonic speed at the points indicated in Fig. 3.1. Beyond this region we expect to see a turbulent boundary layer forming which transports material off the surface. The cloud will compress along the line of motion due to an internal shock wave generated by the external gas. From Bernoulli's theorem we know that the pressure is low on the lateral sides which causes an overspilling of the cloud due to the high inner pressure of the compressed cloud (Doroshkevich & Zeldovich, 1981). This causes mass loss irrespective of any instability.

The Kelvin-Helmholtz Instability — Kelvin-Helmholtz instabilities (KHI) occur when velocity shear is present at the interface between two fluids. The importance of the KHI, in the context of gas cloud stability, has been studied by many authors e.g. Nulsen (1982), Murray et al. (1993), Vietri et al. (1997), Mori & Burkert (2000).

Neglecting gravity, the dispersion relation of the KHI, in the notation of our setup, for an incompressible fluid is (Chandrasekhar, 1961)

$$w = k\frac{(\rho_2\rho_{\rm cl})^{1/2}v_2}{(\rho_2+\rho_{\rm cl})} \approx \frac{kv_2}{\chi^{1/2}}, \tag{3.7}$$

where k is the wavenumber of the instability and the last approximation holds for $\chi \gg 1$. The characteristic growth time for the KHI is then

$$\tau_{\rm KH} \equiv \frac{2\pi}{w} = \frac{2\pi(\rho_2+\rho_{\rm cl})}{k(\rho_2\rho_{\rm cl})^{1/2}v_2} \approx \frac{2\pi\chi^{1/2}}{kv_2}. \tag{3.8}$$

By naively using the post-shock quantities of Eq. 3.1 and our choice of cloud parameters, we can calculate an approximate time dependence of the KH instability, which is shown in Fig. 3.3, (blue, solid lines) for perturbations of size $R_{\rm cl}$ (thick), $R_{\rm cl}/2$ (middle) and $R_{\rm cl}/3$ (thin). Small scale instabilities grow faster due to the $\tau_{\rm KH} \sim k^{-1}$ relation. The first modes to grow are the shortest. Their growth will act to widen the interface between the shearing layers, hence dampening the growth of modes smaller than the thickness of the interface (Chandrasekhar, 1961). The fastest growing modes are now those that are equal to the thickness of the interface. As this process continues, the mode responsible for the cloud destruction is that which is comparable to the size of the cloud itself: $k_{\rm cl} \sim 2\pi/R_{\rm cl}$ (Nulsen, 1982; Murray et al., 1993).

The instability growth time is always larger than the cloud crushing time. The horizontal line at $\tau = 1.6\,\tau_{\rm cr}$ in Fig. 3.3 indicates roughly the time at which the $k_{\rm cl}$ KH mode should have grown fully. We will from now on refer to this time as $\tau_{\rm KH}$.

Note that cloud compressibility can be taken into account when calculating the KH growth time (see Vikhlinin et al., 2001), but was omitted for simplicity. Also note that in certain more physically motivated situations with external gravitational fields, self-gravity, physical viscosity, magnetic fields, radiation etc., the KHI is modified and is damped in most but not all cases (e.g. Murray et al., 1993; Vietri et al., 1997; Miniati et al., 1999; Gregori et al., 2000).

The Rayleigh-Taylor Instability — Rayleigh-Taylor instabilities (RTI) occur when a denser fluid is accelerated by a less dense fluid. The cloud is accelerated with respect to the background and we expect RTIs to develop. The dispersion relation for the RTI is (Chandrasekhar, 1961)

$$|w^2| = k'a\left(\frac{\rho_{\rm cl}-\rho_{\rm ext}}{\rho_{\rm cl}+\rho_{\rm ext}}\right) \approx k'a, \tag{3.9}$$

where the last approximation is valid for $\chi \gg 1$. The KHI, which results from shearing flows, has a 2D geometry, and can be described by as single wave vector k. By contrast, the RTI necessarily has a 3D geometry and must be described by a *vector* wavelength, $\underline{k}' = (k_1, k_2)$, of magnitude: $k' = \sqrt{k_1^2 + k_2^2}$. The acceleration on the surface can be assumed to be $a = \epsilon a_{\rm cl}$, where $a_{\rm cl}$ is given by Eq. 3.3 and ϵ is an efficiency factor. Note that it is very difficult to analytically determine the efficiency of the momentum transfer from the external medium onto the cloud. By using $\epsilon = 1$ we will get a lower limit on $\tau_{\rm RT}$.

Fig. 3.3 shows, for our choice of parameters, the characteristic growth times for RT instabilities (red, dashed lines) of size $R_{\rm cl}$ (thick), $R_{\rm cl}/2$ (middle) and $R_{\rm cl}/3$ (thin), demonstrating that $\tau_{\rm KH} < \tau_{\rm RT}$ for large instabilities. The largest mode grows very slowly and is probably not important in this type of problem. However, we expect that a fast growing small-scale RT instability should develop on the cloud front, especially on the axis of symmetry as the flow rams into the stagnation point. Complicated mixtures of KHIs and RTIs during later evolution is also expected until the cloud becomes fully co-moving with the flow.

3.4. Numerical Simulations

Our numerical simulations solve the Euler equations which neglect physical viscosity and radiative processes; we assume a perfect gas equation of state $P = R\rho T/M$ where R is the gas constant and M is the molar mass. Away from shocks, the evolution is strictly adiabatic. This means the gas can only undergo reversible heating and cooling by adiabatic compression or expansion, or irreversible heating in shocks. In order to isolate the differences in hydrodynamic solvers, we neglect the self-gravity of the gas.

Initial Conditions — The initial conditions (IC) for the blob test are set up in the following way: we use a periodic simulation box of size, in units of the cloud radius $R_{\rm cl}$, $\{L_x, L_y, L_z\} = \{10, 10, 40\}$ and we center the cloud at $\{x, y, z\} = \{5, 5, 5\}$. The ICs are generated by randomly placing equal mass particles to obtain the correct densities and cloud radius. Using an SPH code, the system is evolved and allowed to relax to obtain pressure equilibrium. By repeatedly adding small random velocities to the particles and letting the system relax we obtain a glass-like IC. Random velocities from spurious pressure forces will in this way be minimized compared to a completely random IC. Once the glass is created, the streaming velocity $v_{\rm ext}$ is given to the particles constituting the hot ambient medium. Due to the glass IC we note that the random velocities coming from spurious pressure forces are $\lesssim$ few % of $v_{\rm ext}$. One could smoothly increase the velocities to be more faithful to astrophysical situations, but this more violent start together with particle noise serves as the initial seed for surface instabilities of the cloud. Formally this can be seen as a triggering of small scale RT and Richtmyer-Meshkov instabilities. The Richtmyer-Meshkov instability occurs when a contact discontinuity gets shocked or rapidly accelerated. This generates vorticity and structures similar to those of RT (e.g. Inogamov, 1999).

This particle setup is used as IC for the SPH simulations. The ICs for the grid simulations are obtained by smoothing the gas quantities (density, temperature and velocities) onto each cell center using the same spline kernel as in the SPH codes (see Section 3.4) using 32 nearest neighbours. In this way we have a consistent setup for both of the methods and the noise introduced by using discrete particles in the SPH simulations is also present in the grid IC. As we will argue below, the key parameters to study are those connected with the resolution and strength of artificial viscosity therefore our parameter space studies will focus on the effect of these.

The Codes — The simulation was carried out with about a dozen different independent simulation codes. Since all the grid codes gave consistent results, and similar for the SPH codes, we shall just present the detailed analysis of a selection of these codes which are summarized in Table 3.1. Here we give a brief description of these codes and the methods used for solving the hydrodynamical equations:

□ ART (AMR) — ART (Adaptive Refinement Tree) is a N-body+gas dynamics AMR code (Kravtsov, 1999; Kravtsov et al., 2002). The ART code uses second-order shock-capturing Godunov-type solver (Colella & Glaz, 1985) to compute numerical fluxes of gas variables through each cell interface, with "left" and "right" states estimated using piecewise linear reconstruction (van Leer, 1979). This is a monotone method that is known to provide good results for a variety of flow regimes and resolves shocks within $\approx 1-2$ cells. A small amount of dissipation in the form of artificial diffusion is added to numerical fluxes (Colella & Woodward, 1984), as is customary in the shock-capturing codes. The details of the flux evaluation and summation on mesh interfaces can be found in Khokhlov (1998). In the simulations presented in this paper, a new distributed MPI version of the ART code developed by Douglas Rudd and Andrey Kravtsov was used (Rudd & Kravtsov, in preparation).

□ CHARM (AMR) — CHARM is an N-body+gas dynamics, AMR code, based on the CHOMBO-AMR library, employing a higher order Godunov's method for the solution of the hydrodynamic equations (Miniati & Colella, 2006). Here a piecewise linear reconstruction scheme with Van Leer's limiter and a nonlinear Riemann solver were used, resulting in a second order accurate method in both space and time. CHARM was used to test the influence of ICs on the cloud evolution.

□ ENZO-PPM (AMR) — ENZO[22] is an Eulerian AMR hybrid code (N-body+gas dynamics) code that was originally written by Greg Bryan and Michael Norman at the National Center for Supercomputing Applications at the University of Illinois (Bryan & Norman, 1997). ENZO uses the Piecewise parabolic method (Colella & Woodward, 1984) for solving fluid equations and has been adapted for cosmology (Bryan et al., 1995). PPM is a higher order accurate version of Godunov's method with an accurate piecewise parabolic interpolation and a non-linear Riemann solver for shock conditions. The method is third order accurate in space and second order in time for fixed timestepping. For variable timestepping it is formally second order in space. This together with the Riemann solver results in a very accurate shock treatment compared to the SPH codes where artificial viscosity is used. In all of our tests we used the Dual Energy Formalism in ENZO. Formally, the use of this is only necessary in hyper-Machian flows ($E_{\rm therm}/E_{\rm tot} \sim 10^{-3}$) to keep the PPM solver stable, and hence makes little difference for our case.

□ ENZO-ZEUS (AMR) — ENZO includes an implementation of the finite-difference hydrodynamic algorithm employed in the compressible magnetohydrodynamics code ZEUS (Stone & Norman, 1992a,b). Fluid transport is solved on a Cartesian grid using the upwind, monotonic advection scheme of van Leer (1977) within a multistep (operator split) solution procedure which is fully explicit in time. This method is formally second order-accurate in space but first order-accurate in time.

The ZEUS method uses a von Neumann-Richtmyer artificial viscosity to smooth shock discontinuities that may appear in fluid flows and can cause a break-down of finite-difference equations. The artificial viscosity term is added in the source terms as

$$\rho\frac{\partial \mathbf{v}}{\partial t} = -\nabla p - \rho\nabla\phi - \nabla\cdot\mathbf{Q} \tag{3.10}$$

$$\frac{\partial e}{\partial t} = -p\nabla\cdot\mathbf{v} - \mathbf{Q}:\nabla\mathbf{v}, \tag{3.11}$$

where $\mathbf{v}$ is the baryon velocity, ρ is the mass density, p is pressure, e is internal energy density of gas and $\mathbf{Q}$ is the artificial viscosity stress tensor, such that:

$$Q_{ii} = \begin{cases} Q_{\rm AV}\rho(\Delta v_i)^2, & \text{for } \Delta v_i < 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.12}$$

and

$$Q_{ij} = 0 \quad \text{for} \quad i \neq j. \tag{3.13}$$

Δx_i and Δv_i refer to the width of the grid cell along the i-th axis and the corresponding difference in gas velocities across the grid cell, respectively. $Q_{\rm AV}$ is a constant that roughly tells us over how many grid zones we smooth shocks. While the correct Rankine-Hugoniot jump conditions are achieved, shocks are thus not treated as true discontinuities. This may cause unphysical pre-heating of gas upstream of the shock wave, as discussed in e.g. Anninos & Norman (1994) and O'Shea et al. (2005).

□ FLASH (AMR) — FLASH[23] is an AMR hybrid code (N-body+gas dynamics) developed by the ASC Center at the University of Chicago (Fryxell et al., 2000). The PPM hydrodynamical solver is formally accurate to second order in both space and time but performs the most critical steps to third- or fourth-order accuracy. For the simulations performed in this paper we have used the publicly available FLASH version 2.3 using AMR with maximum refinement up to the resolutions indicated in Table 3.1.

□ GASOLINE (SPH) — GASOLINE is a parallel Tree + SPH code, described in Wadsley et al. (2004). The code is an extension to the N-Body gravity code PKDGRAV developed by Stadel (2001). GASOLINE uses artificial viscosity (AV) to resolve shocks and has an implementation of the shear reduced version (Balsara, 1995) of the standard (Monaghan, 1992) artificial viscosity. GASOLINE solves the energy equation using the asymmetric form and conserves entropy closely. It uses a standard spline smoothing kernel (Monaghan, 1992) with compact support for the softening of the gravitational and SPH quantities. The kernel is symmetrized by using kernel averaging (Hernquist & Katz, 1989) and we smooth over the 32 nearest neighbours when estimating fluid quantities.

---

[22] ENZO is available at http://lca.ucsd.edu/portal/software/enzo

[23] FLASH is available at http://flash.uchicago.edu/website/home/

The AV is implemented by solving a momentum equation of the form

$$\frac{d\vec{v}_i}{dt} = -\sum_{j=1}^{n} m_j \left(\frac{P_i}{\rho_i^2} + \frac{P_j}{\rho_j^2} + \Pi_{ij} \right) \nabla_i W_{ij}, \tag{3.14}$$

where P_j is pressure, $\vec{v}_i$ is velocity, W_{ij} is the smoothing kernel and the AV term Π_{ij} is given by

$$\Pi_{ij} = \begin{cases} \frac{-\alpha\frac{1}{2}(c_i+c_j)\mu_{ij}+\beta\mu_{ij}^2}{\frac{1}{2}(\rho_i+\rho j)} & \text{for } \vec{v}_{ij}\cdot\vec{r}_{ij} < 0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.15}$$

$$\text{where } \mu_{ij} = \frac{h(\vec{v}_{ij}\cdot\vec{r}_{ij})}{\vec{r}_{ij}^{\,2} + 0.01(h_i+h_j)^2}, \tag{3.16}$$

where $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$, $\vec{v}_{ij} = \vec{v}_i - \vec{v}_j$ and c_j is the sound speed. α and β are the coefficients used for setting the viscosity strength, and are essential for capturing shocks and preventing particle interpenetration. Note that the viscosity term vanishes for non-approaching particles and the β-parameter is the SPH implementation of the Neumann-Richtmeyer artificial viscosity. The commonly used values in the literature is $\alpha = 1$ and $\beta = 2$ which originally was proposed by Lattanzio et al. (1986) using Sod shock tube tests. Later we will carry out experiments with different values of α and β.

☐ GADGET-2 (SPH) — The TreeSPH code GADGET-2[24] (Springel, Yoshida & White, 2001; Springel, 2005) is the updated version of the GADGET-1. The code is similar in character to GASOLINE but uses an entropy conserving formulation of SPH. This means that the thermodynamic state of each fluid element in GADGET-2 is defined through the specific entropy and not the specific thermal energy. GADGET-2 uses a somewhat different formulation of artificial viscosity than GASOLINE. The viscosity term in Eq. 3.14 is here formulated as

$$\Pi_{ij} = -\frac{\alpha}{2}\frac{v_{ij}^{sig} w_{ij}}{\rho_{ij}}, \tag{3.17}$$

where $v_{ij}^{sig} = c_i + c_j - 3w_{ij}$ is the so called signal velocity. Here $w_{ij} = \vec{v}_{ij} \cdot \vec{r}_{ij}/|\vec{r}_{ij}|$ is the relative velocity projected onto the separation vector provided particles approach each other. Like GASOLINE, GADGET uses a spline smoothing kernel (Monaghan, 1992) and we employ smoothing over the 32 nearest neighbours. In our test we used the publicly available GADGET-2 version 2.01.

3.5. Results

Fig. 4 shows central density slices of GASOLINE (Gas_10m), GADGET-2 (Gad_10m), ENZO (Enzo_256), FLASH (FLASH_256) and ART (ART_256). These are the high resolution simulations with the default standard settings.

The simulations of the two SPH codes, Gas_10m and Gad_10m, show a very similar evolution. As expected, a detached bow shock forms directly in front of the cloud. An internal shock wave forms within the cloud compressing it. The post shock flow encompasses the cloud, creating Bernoulli zones on the top and bottom with lower pressure. This causes the cloud to become elongated as well as compressed along the z-axis and we see gas being ablated, i.e. stripped through the induced pressure differences, from the top and bottom edges. Gas stripping slowly progresses and the cloud's shape does not change significantly for a long time. Fig. 3.5 shows the particles in a thin slice centered on the cloud. The velocity vectors of each particle are plotted in a reference frame centered on the cloud. The colours indicate the gas density. Behind the edges of the cloud we see a vortex created due to the shearing motion of the ambient medium which creates a low pressure region behind the cloud.

Initially, the cloud evolution is similar in the grid simulations. It is compressed and elongated and gas is removed from the trailing edges where the vortex has created a vacuum behind the cloud. Some of the ambient medium is entrained in the turbulent wake behind the cloud and falls onto the backside of the cloud. However, the late cloud evolution is very different in these simulations. Early on we observe surface perturbations

---

[24] GADGET-2 is available at http://www.mpa-garching.mpg.de/gadget/

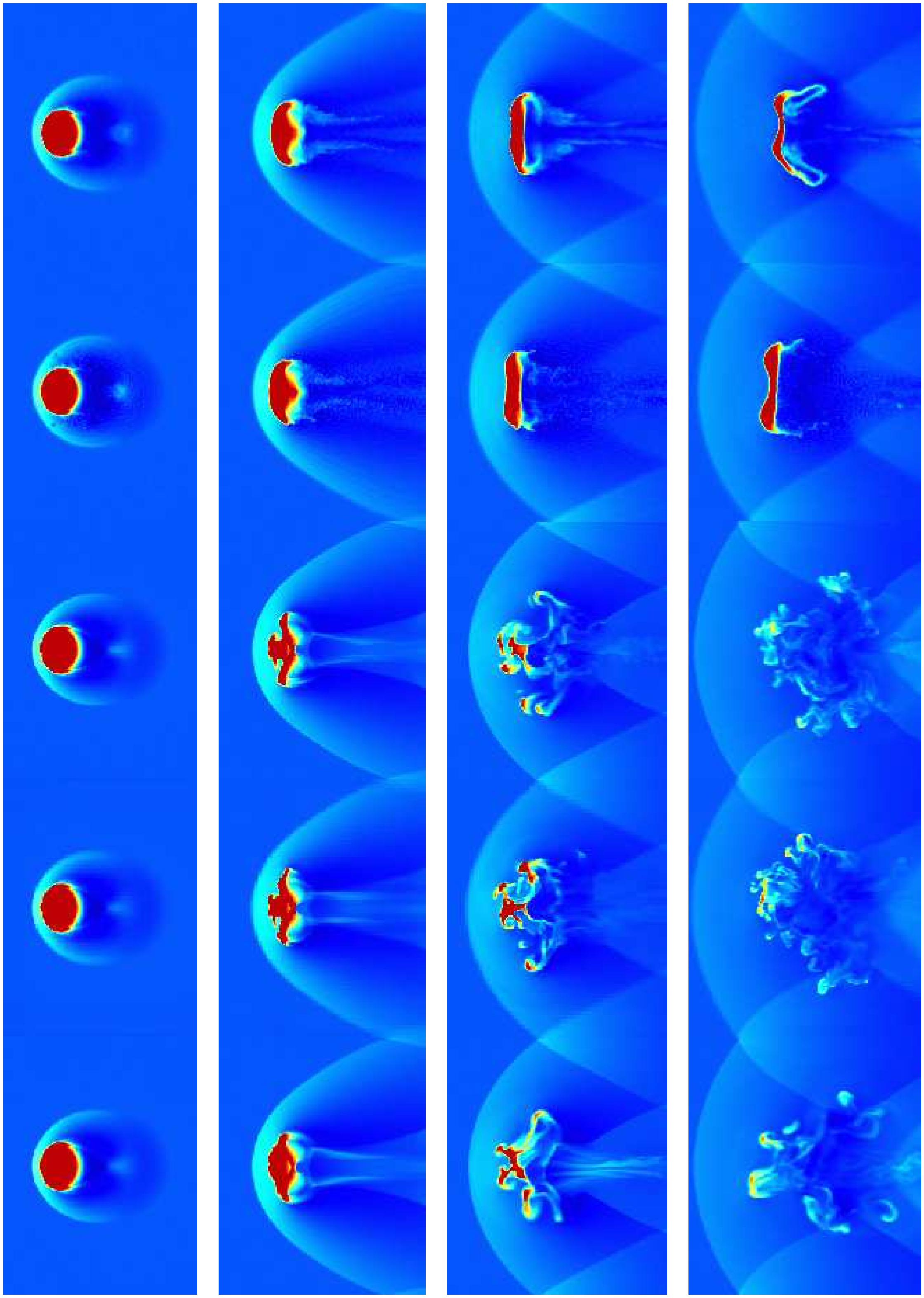

Figure 3.4: Gas density slices through the center of the cloud at $t = 0.25, 1.0, 1.75$ and $2.5\,\tau_{\rm KH}$. From top to bottom we show GASOLINE (Gas_10m), GADGET-2 (Gad_10m), ENZO (Enzo_256), FLASH (FLASH_256) and ART-Hydro (ART_256). The grid simulations clearly show dynamical instabilities and complete fragmentation after $2.5\,\tau_{\rm KH}$, unlike the SPH simulations in which most of the gas remains in a single cold dense blob.

Table 3.1: Simulation details. ENZO and ART use the static grids indicated in the table while the CHARM and FLASH simulations have been run using AMR up to the indicated resolution. All static grid as well as the FLASH simulations were initialized using the stated resolution, CHARM started from 32,32,128.

nParticles/Grid size	AV	Name
ART, static		
64,64,256	no AV	ART_64
128,128,512	no AV	ART_128
256,256,1024	no AV	ART_256
CHARM, AMR		
512,512,2048	no AV	CHARM_512
ENZO-PPM, static		
64,64,256	no AV	Enzo_64
128,128,512	no AV	Enzo_128
256,256,1024	no AV	Enzo_256
ENZO-ZEUS, static		
256,256,1024	$Q_{\rm AV} = 2.0$	Enzo_ZEUS1
256,256,1024	$Q_{\rm AV} = 0.5$	Enzo_ZEUS2
256,256,1024	$Q_{\rm AV} = 0.1$	Enzo_ZEUS3
FLASH, AMR		
64,64,256	no AV	FLASH_64
128,128,512	no AV	FLASH_128
256,256,1024	no AV	FLASH_256
GADGET-2		
10^7	$\alpha = 0.8$	Gad_10m
GASOLINE		
10^6	$\alpha = 1.0$, $\beta = 2.0$	Gas_1m
10^7	$\alpha = 1.0$, $\beta = 2.0$	Gas_10m
10^7	$\alpha = 0$, $\beta = 2.0$	Gas_10mAV1
10^7	$\alpha = 0$, $\beta = 0.5$	Gas_10mAV2
10^7	$\alpha = 0$, $\beta = 0.1$	Gas_10mAV3
10^7	Balsara, $\alpha = 1.0$, $\beta = 2.0$	Gas_Bals

on the front of the cloud, probably originating from the way the ICs are setup (see argument in Section 3.4). A complicated mixture of KHIs and RTIs are developing on the cloud front which, due to subsequent compression and lateral expansion, becomes even more KH and RT unstable. By $t \sim \tau_{\rm KH}$, large scale KHIs have developed and the cloud starts to fragment. Further instabilities and turbulence mixes the smaller clumps of gas into the ambient medium. All grid simulations show basically the same cloud destruction time. We also note that Eulerian (shock capturing) methods effectively localize shocks to a few grid cells compared to the smoothed out shocks in the SPH simulations resulting from AV shock capturing schemes.

In Fig. 3.6 we show the remaining cloud mass fractions as a function of time for the ENZO and GASOLINE simulations. These are representative of grid and SPH methods. We define the cloud as being any gas that satisfies $T < 0.9\,T_{\rm ext}$ and $\rho > 0.64\,\rho_{\rm cl}$. It is of course possible to construct more elaborate criteria but these select the gas that visually is a part of the cloud. The figure shows that both techniques give a similar mass loss up to $\sim \tau_{\rm KH}$. Before this time the gas loss is mainly due to ablation into the low pressure zone created behind the cloud. As soon as we pass $\tau_{\rm KH}$ for large scale KHIs the SPH and grid methods diverge. In the grid simulation, the cloud quickly disrupts and diffuses into the ambient medium, while the SPH simulation only

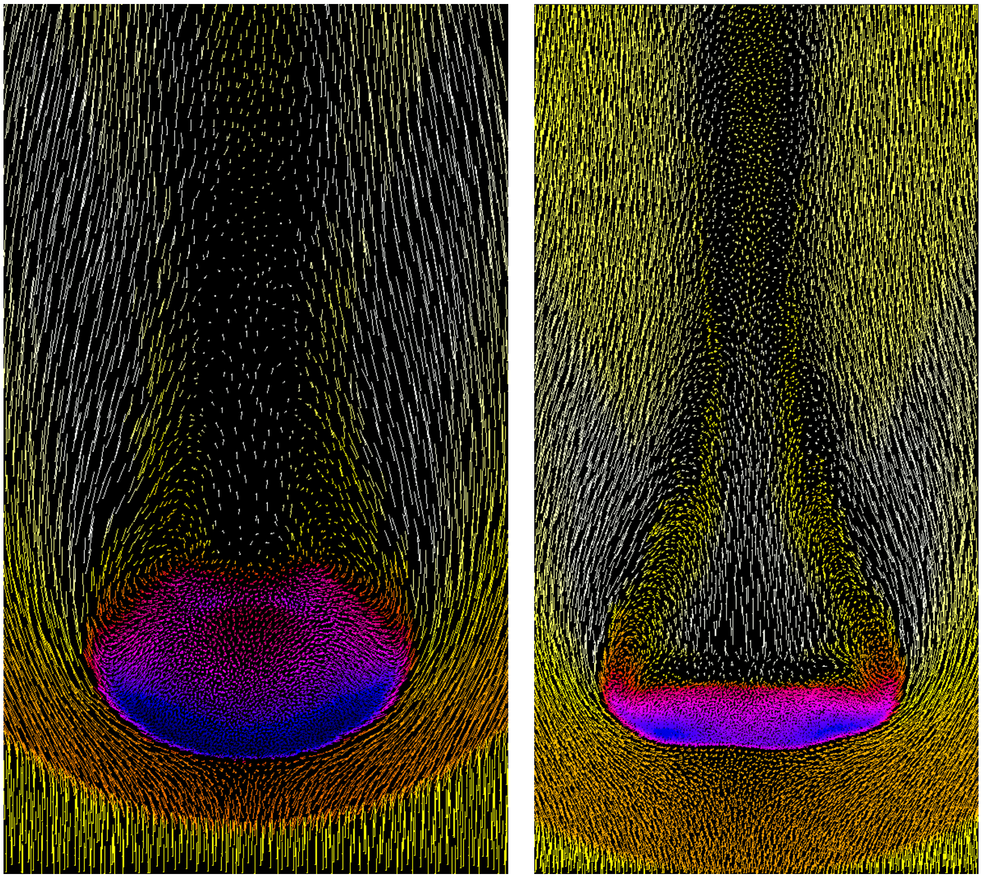

Figure 3.5: A thin central slice of the SPH particles of Gas_10m at $t = 0.75\,\tau_{\rm KH}$ (left panel) and $t = 1.5\,\tau_{\rm KH}$ (right panel). The density ranges from high (blue) to low (white) and the magnitude of the velocity vectors are normalized to the reference frame of the center of the cloud. We clearly see the effect of the cloud stretching due to the lateral Bernoulli zones and the formation of downstream vorticity.

shows continuing stripping. After $t = 2.5\,\tau_{\rm KH}$, no gas in the grid simulation can satisfy our criteria while the SPH simulation still shows a mass fraction $\approx 40\%$. This shows us that the vortex shedding through the Bernoulli zones is the most important mechanism for mass-loss at $t < \tau_{\rm KH}$ in both methods. After this time dynamical instabilities dominate the grid mass-loss.

Resolution Dependence — It is difficult to do a direct translation between grid and SPH resolution. The maximum allowed resolution is a fixed grid of size $256 \times 256 \times 1024$ in the grid runs and 10^7 particles in the SPH runs. This means that there is almost a factor 7 more cells compared to particles. On the other hand, cells are uniformly distributed in space and only $\approx 70\,276$ cover the cloud in the initial setup of an almost perfectly spherical cloud. This should be compared to the $\approx 10^5$ particles constituting the cloud in the high resolution SPH run. A comparison like this is still not straightforward due to the fact that SPH uses particles as *non-independent* resolution elements. This means that each particle is not a carrier of information without neighbours to smooth over, and the effective number of resolution elements is more or less set by the kernel shape and number of neighbours to smooth over.

Resolution affects the convergence of hydrodynamical simulations. A cut-off is always introduced on the scale of the spatial resolution below which instabilities can not be resolved. This often serves as a source of numerical viscosity. For most of the codes used in the comparison, we have varied the resolution in order to obtain an understanding of how this changes the cloud morphology, mass loss and fragmentation time (see Table 3.1). Fig. 3.7 shows the outcome of, from top to bottom, Enzo_64, Enzo_128, Enzo_256,

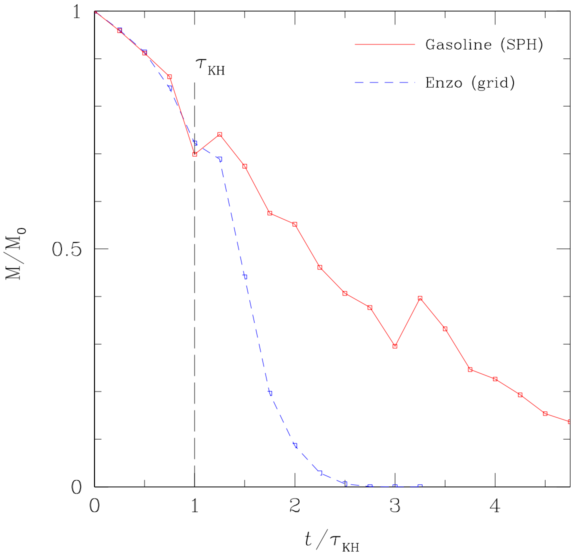

Figure 3.6: The evolution of the cloud mass fraction. In the SPH simulation (solid, red), the cloud slowly loses mass to the ambient medium and has not been completely mixed even after $5\,\tau_{\mathrm{KH}}$. The grid simulation (dashed, blue) follows the SPH up to the time at which the KH instability causes it to rapidly fragment and mix.

Gas_1m, and Gas_10m. In the grid simulations we conclude that, while the compression and elongation of the cloud is relatively similar, the detailed way the cloud fragments is resolution dependent as the IC's also are (Jones et al., 1996). In Enzo_64, a mode of the KHI symmetric with respect to the symmetry axis of the cloud is dominant. This mode becomes less dominant as resolution is increased (Enzo_128) and has not yet developed by $t = 1.5\,\tau_{\mathrm{KH}}$ in Enzo_256. Going to higher resolution we see more and more small scale instabilities developing which enhance mixing of the cloud material with the background flow. Numerical diffusion is stronger in low resolution simulations which is why parts of the cloud survive longer in the higher resolution runs. The different SPH simulations are qualitatively very similar. Instabilities can not be resolved in Gas_1m nor in Gas_10m. However, we note a weak large scale RTI on the cloud front at $t = 2.25\,\tau_{\mathrm{KH}}$ in Gas_10m, which is absent in Gas_1m.

The general description above is again quantified by studying the cloud mass fraction at each time step, see Fig. 3.8. In this plot we have also added an extra low resolution SPH simulation using only 10^5 particles. The grid simulations show a clear trend of dissolving the cloud quickly after $\sim \tau_{\mathrm{KH}}$ regardless of resolution while the SPH simulations only show a steady mass loss due to the material ablated into the trailing vacuum. Decreasing the SPH resolution causes the mass fraction to rise above the initial value during the initial phase and mass is lost more rapidly for $t > \tau_{\mathrm{KH}}$. The latter effect is most probably due to the increased mass of each particle, causing each particle interaction to transfer momentum in a more violent, “bullet-like” fashion.

SPH vs. Grid Resolution Criteria — In the study of Mac Low & Zahnle (1994), simulations of the impact of comet Shoemaker-Levy with Jupiter were carried out using the ZEUZ-2D grid code (Stone & Norman, 1992a). They found that a minimum of 25 grid cells per cloud radius were required to follow the evolution correctly. This resolution is reached in all of the high resolution grid simulation that we have performed, but it is very important to note that the destruction of the cloud is captured even in the lowest resolution runs, where we only have 7 cells per radius. In the case of SPH it is, as mentioned in Section 3.5, more difficult to apply this criterion. The most conservative translation of the criterion is to use 25 non-overlapping smoothing kernels per radius. This is indeed a lower limit to the resolution as the cubic spline smoothing kernel used in our SPH tests can not exactly be interpreted as a grid cell. The kernel is given a radius allowing it to encompass 32 particles, and the strength of the kernel falls off rapidly. At half the kernel radius (at h), only $1/4$ of the kernel's central value remains (Monaghan, 1992), indicating that we probably have more resolution elements than in the non-overlapping kernel case. To safely test the resolution criterion, we anyway adopt the conservative 25 independent kernel interpretation in this section.

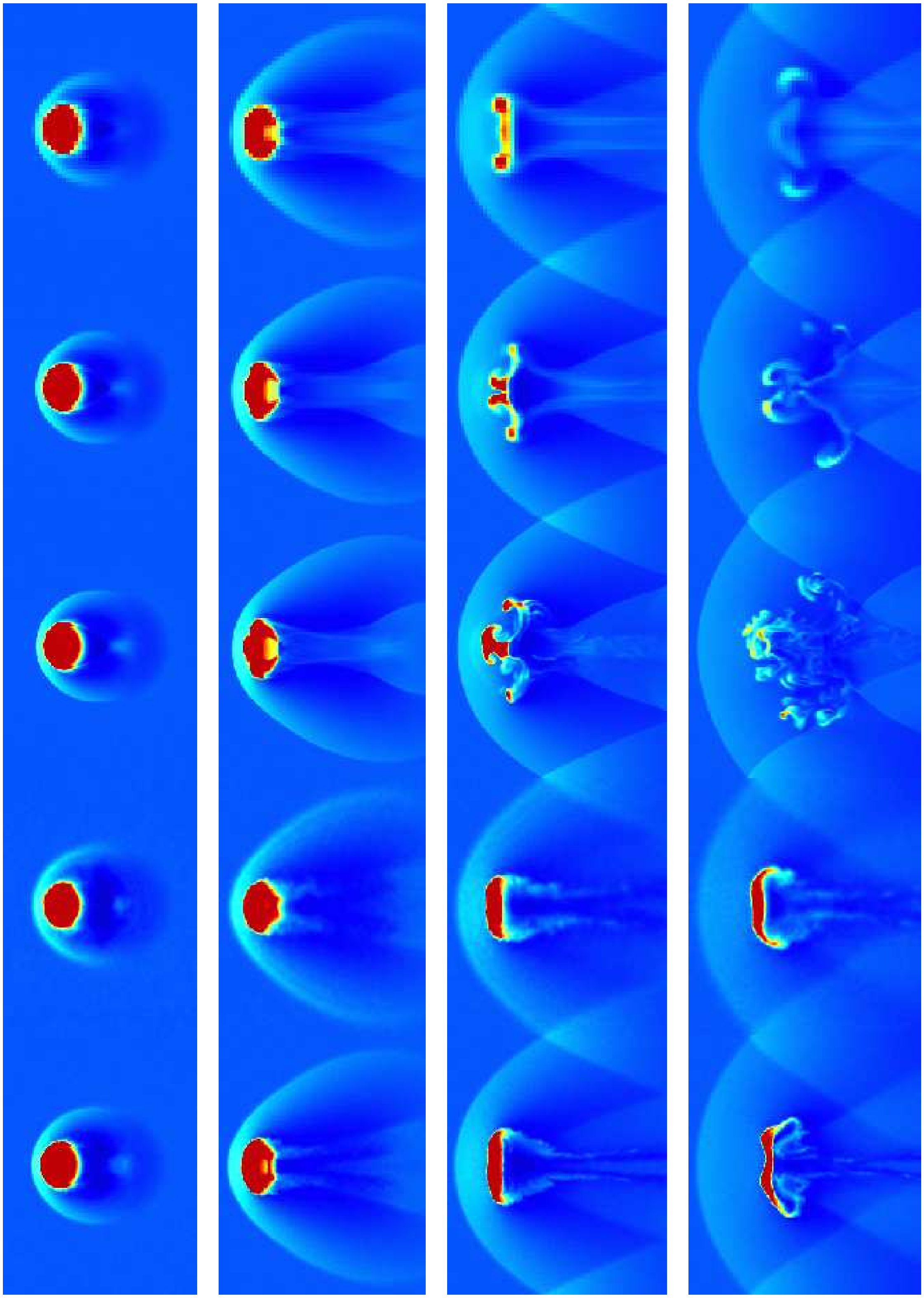

Figure 3.7: Resolution study for ENZO and GASOLINE. The panels show density slices of, from top to bottom, Enzo_64, Enzo_128, Enzo_256, Gas_1m and Gas_10m for $t = 0.25, 0.75, 1.5$ and $2.25\,\tau_{\mathrm{KH}}$. We see that resolution changes the phase of the instabilities in the grid simulations while the destruction time is the same. Higher resolution also shows less diffusion and better resolves small scale fragments. The GASOLINE runs are not able to resolve small scale instabilities at all.

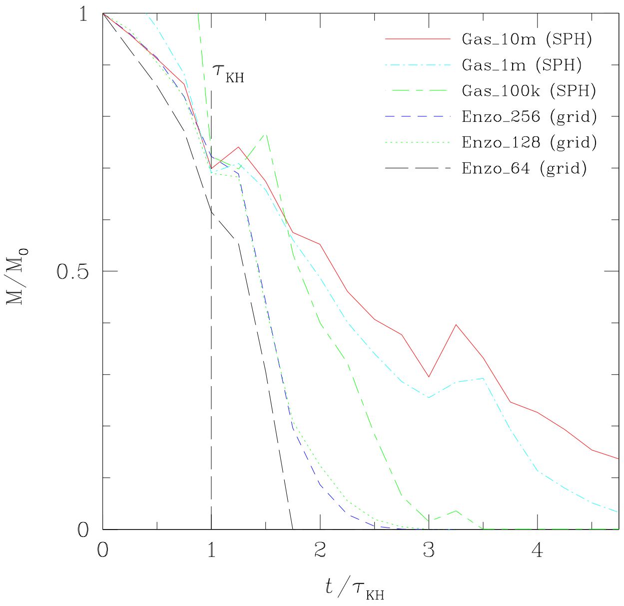

Figure 3.8: The evolution of the cloud mass fraction for different resolutions. As the resolution of the grid simulations is increased from 64 to 128 to 256 cells across the wind tunnel, the amount of mass increases a little but converges. Increasing the resolution of the SPH simulations does not decrease the amount of mass lost, rather the opposite, perhaps due to the momentum transfer due to massive particles acting like "‘bullets"’.

In order to investigate this we perform two additional simulations using ENZO and GASOLINE. Any uniform spherical distribution of N_p particles using a kernel smoothing over the n nearest neighbours has $n_\mathrm{k} = \frac{1}{2}\left(\frac{N_\mathrm{p}}{n}\right)^{1/3}$ independent smoothing kernels covering one radius. By having $n_\mathrm{k} = 25$ and $n = 32$ we see that we require $N_\mathrm{p} = 4 \times 10^6$ *in the cloud only*. This is to be compared with the requirement using grid codes which is $\approx 65\,450$ cells, a substantial difference in computation and storage. We use a smaller box in order to manage the large simulation required for SPH. In units of the cloud radius R_cl, the sides are $\{L_x, L_y, L_z\} = \{4, 4, 12\}$ where we center the cloud at $\{x, y, z\} = \{2, 2, 3\}$. The smaller box will give us a setup that is not “as clean” as the previous ones as the backflow and lateral bow shock interacts with the cloud in the later evolution due to the periodic boundary condition that are necessary to impose (inflowing boundary conditions are not possible in the current version of GASOLINE). We will however fully trace the evolution past the important τ_KH, which is estimated in the same way as in Section 3.3. To facilitate computations we used a density contrast $\chi = 20$ for this test. This reduces the total number of SPH particles to $\approx 1.36 \times 10^7$. To optimize the conditions for the SPH simulations we have adopted a lower viscosity setting than normally used; $\alpha = 0.1$ and $\beta = 1.5$. This is done in order not to suppress possible growth of instabilities while still capturing shocks (see Section 3.6 for a discussion). The grid simulation is performed as before but now using a static grid of size $100 \times 100 \times 300$. In order to see the direct effect of the high resolution run we have also performed a simulation using the same number density of particles in the cloud as in Gas_10m but with the new density contrast, viscosity setting and box size. The visual outcome of the simulations can be seen in Fig. 3.9. The conclusion of the previous sections remains valid; the initial phase of the evolution is very similar for the grid and SPH simulations. However, later evolution of the cloud in the grid simulations shows surface instabilities developing leading to fragmentation and mixing of material after $t = \tau_\mathrm{KH}$. The cloud in the SPH simulation does not fragment and suffers only from lateral elongation and ablation.

The differences are small between the high resolution and standard SPH simulation with only minor morphological differences probably owing to different capturing of the more complicated shock structure in this new setup. A test of the standard resolution simulation using larger viscosity setting was also performed (not shown here) which produces identical results, assuring us that the specific viscosity setting is not unphysically low.

We conclude that the observed differences between grid and SPH methods are not related to resolution and

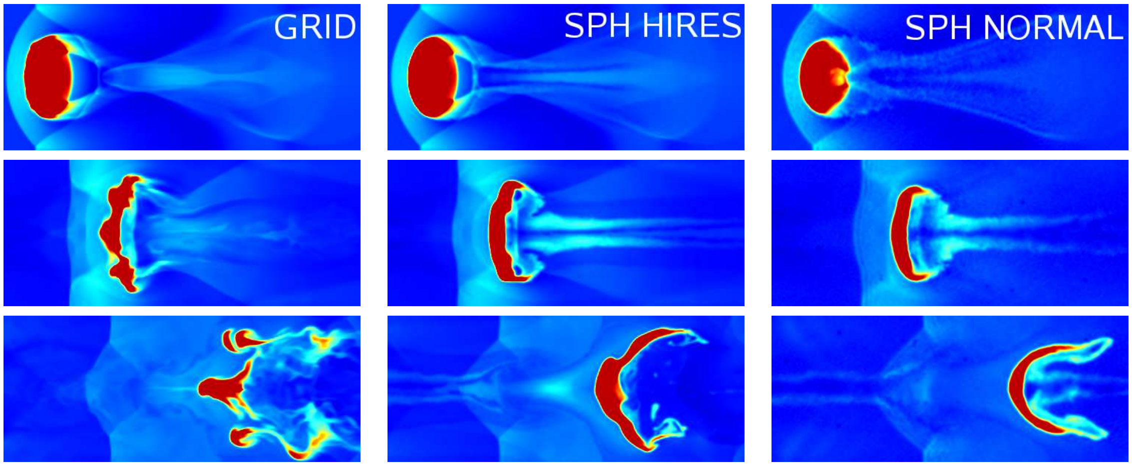

Figure 3.9: Each frame shows a density slice through the cloud center at times $t = 0.5, 1.0$ and $1.5\,\tau_{\rm KH}$ with densities varying from low (blue) to high (red). The grid (ENZO) simulation (left) shows instabilities developing on the surface causing the cloud to fragment, while these features are absent in the SPH (GASOLINE) simulation (middle and right).

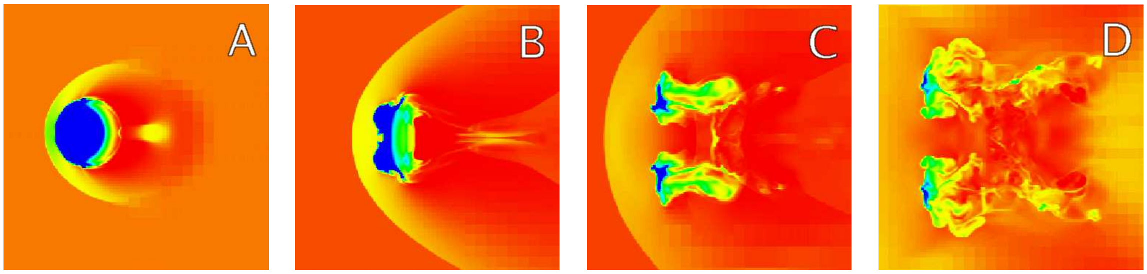

Figure 3.10: Evolution of the cloud with 'analytic' initial conditions using the CHARM code. Each frame shows a density slice through the cloud center at times $t = 0.24, 0.9, 1.7$ and $2.5\,\tau_{\rm KH}$ with densities varying from low (red) to high (blue).

that convergence must be reached by other means.

Initial Seeds — As partly shown in the previous test, the development of the instabilities, particularly during the nonlinear stages, is sensitive to the exact definition of the initial conditions. This is because they set the seed perturbations out of which the instabilities grow. However, while the mixing of the cloud material with the background medium is affected by small scale motions that arise from the small unstable scales, the cloud disruption is mostly the result of the development of the large scale perturbations. As an example of this in Fig. 3.10 we show the evolution of the cloud-wind interaction but with initial conditions set directly from the analytic definition. Thus in this case the initial conditions are free of noise and are purely symmetric. A base grid of $(32 \times 32 \times 128)$ was used with two additional levels of refinement with refinement ratio of 4 placed dynamically in regions where the relative change in density, $\Delta\rho/\rho$ exceeded 20%. This corresponds to an effective resolution of $512 \times 512 \times 2048$ in the finest grids, which reduces the level of perturbation with respect to the previous cases.

As shown in panel B of Fig. 3.10 the most destructive mode has a different phase than in the cases illustrated above for the corresponding grid based codes. However, as in the previous cases, by $t = 2.5\,\tau_{\rm KH}$ (panel D) the cloud has been completely reduced to debris by the instabilities. This shows that despite differences in the appearance of the cloud gas distribution its fundamental fate of disruption and subsequent mixing on a timescale of a few $\tau_{\rm KH}$ is independent of the specific definition of the initial conditions.

3.6. Why so different?

What is the reason for the observed discrepancies between simulations carried out using SPH and grid-based techniques? Differences between SPH and grid-based results have been discussed before in the literature (Frenk & et al, 1999; Pearce et al., 1999; Thacker et al., 2000; Ritchie & Thomas, 2001; Tittley et al., 2001; Springel & Hernquist, 2002; Marri & White, 2003; O'Shea et al., 2005) in different contexts to this study. While artificial viscosity is the most obvious focus for criticism of SPH it is not the main reason for the differences observed in this test. We will show this in Section 3.6 before focusing on the almost complete suppression of KH (and RT) instabilities in SPH simulations of this test and present an explanation of why this occurs.

Artificial Viscosity — The artificial viscosity β parameter in Eq. 3.15 is necessary for shock capturing and is required for SPH to work properly in unsmooth supersonic flows. In smooth flows where interparticle velocities are vanishing, no AV is required and is turned off (see Eq. 3.15). The α parameter has a less obvious meaning and the classical $\alpha = 1.0$ setting is most probably unphysical. It can be argued (e.g. Watkins et al., 1996) that α can roughly be interpreted as a Navier-Stokes shear plus bulk viscosity, even though the AV is only sensitive to flow properties such as inter-particle traveling. Bulk viscosity is normally not important in fluid dynamics, except in the theory of attenuation of sound waves (e.g. Faber, 1995). In numerical simulation its inclusion is for the most part to dampen the so called *post-shock ringing*. Many grid based techniques employ AV in order to stabilize the solutions from high frequency oscillations occurring at sharp transitions in flux quantities. All of the grid methods in this paper, except for ENZO_ZEUS, use of Godunov's method. This means, among other things, that fluxes are calculated using Riemann solvers, hence not needing any inclusion of explicit AV terms except for very high Mach number shocks (Colella & Woodward, 1984). Note however that there is always numerical viscosity due to resolution and truncation error in *all* simulation methods. In practice, this means that dynamics on the resolution scale is damped. To quantify the effect of this in the form of an effective viscosity term is not straightforward and requires well defined problems with analytically known solutions to test against. The inclusion of AV leads us to one of the first possibilities for the observed discrepancy: *We are not solving the same hydro-dynamical equations in the different codes*. By adding AV we are solving some kind of Navier-Stokes equation when we actually want to compare the solutions to the grid codes that, in this sense, are closer to the Euler equations.

Viscosity has two major effects on the processes we want to capture in this test:

1 – Dampening of small scale velocity perturbations and random velocities.
2 – Diffusion of post shock vorticity and smearing of turbulence.

The effect of (i) will enter as a stabilizing factor for the growth of instabilities. Physical kinematic viscosity, ν, sets a cut off for the size of the smallest eddies in turbulence (Shu, 1992), below which turbulent motion is diffused. The effect of (ii) follows from the first one and is obvious from inspection of the vorticity transport equation (e.g. Shu, 1992)

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + \nabla\times(\boldsymbol{\omega}\times\boldsymbol{v}) = \nabla P\times\nabla\left(\frac{1}{\rho}\right) + \nu\nabla^2\boldsymbol{\omega}\,, \tag{3.18}$$

where $\boldsymbol{\omega} \equiv \nabla\times\boldsymbol{v}$ is the vorticity. The two terms on the right hand side can create or diffuse vorticity. The first term is the *baroclinic* term which is non vanishing if we have non-aligned pressure and density gradients. This is the case in oblique shocks like in the bow shock of our cloud simulation. The second term is responsible for diffusing vorticity in space i.e. taking local vorticity and spreading it into the general flow. This means that as soon as we have viscosity, *we will dampen vorticity*. Especially important is the vorticity generated in the post shock flow, which should act to destabilize the cloud together with the surface instabilities.

A study on how AV dampens small scale vorticity was made by Dolag et al. (2005). By using a low viscosity formulation of SPH they find higher levels of turbulent gas motions in the ICM and noted that shocked clouds tend to be unstable at earlier times. However, by looking at their Figure 3 we note that the overall difference in the cloud evolution is small. As we will see in the tests carried out below, lowering the AV does not necessarily lead to improved results.

In order to understand the effect of artificial viscosity in our cloud-wind test we have performed three simulations with modified setting of the viscosity coefficients. These are Gas_10mAV1, Gas_10mAV2 and

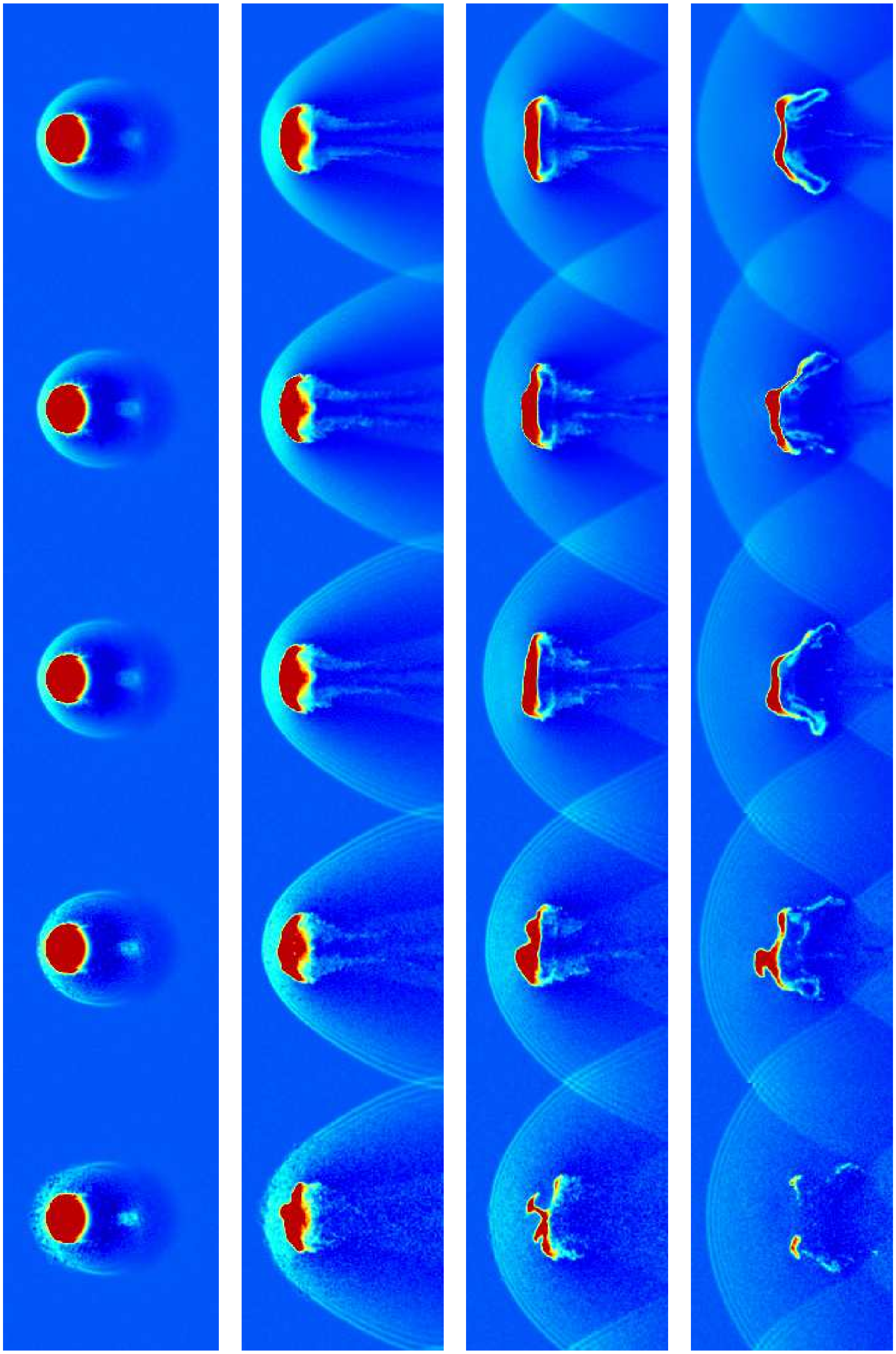

Figure 3.11: Viscosity study for GASOLINE. The panels show density slices of, from top to bottom, Gas_10m, Gas_Bals, Gas_10mAV1, Gas_10mAV2 and Gas_10mAV3 for $t = 0.25, 0.75, 1.5$ and $2.25\,\tau_{\mathrm{KH}}$. We can see how reducing shear viscosity and removing the bulk viscosity renders very similar results; the cloud destabilizes to a higher degree. By reducing the shock capturing viscosity the cloud destabilizes even further, most probably to an unphysical solution in the lower setting. The artificial post shock ringings also gets more pronounced, as expected for lower viscosity settings.

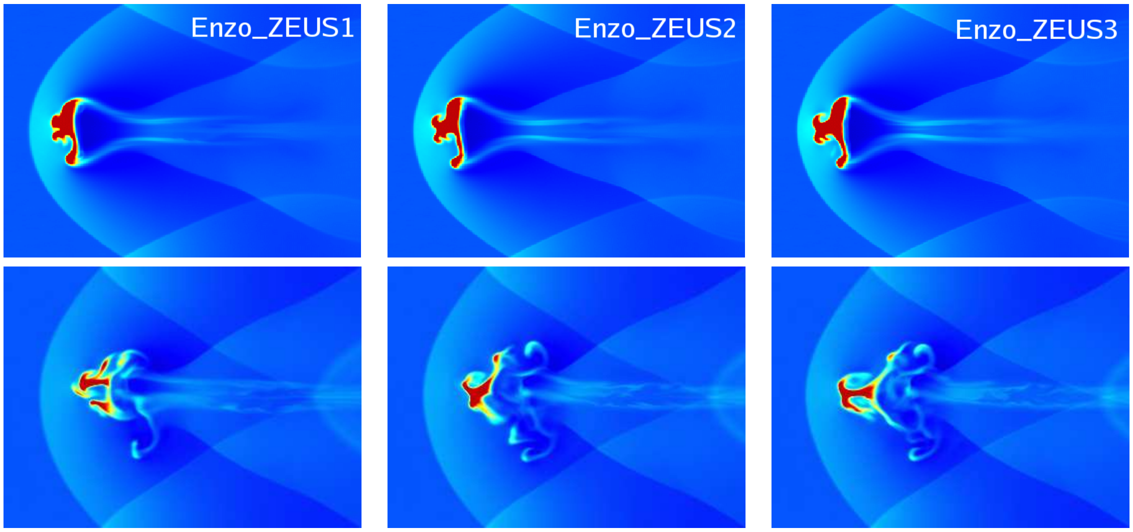

Figure 3.12: Viscosity study for ENZO-ZEUS. The panels show density slices of, left to right, Enzo_ZEUS1, Enzo_ZEUS2 and Enzo_ZEUS3 at $t = 1.5$ and $2.25\,\tau_{\rm KH}$. The outcome of the simulations show little difference; only minor morphological changes are found. Artificial viscosity in ENZO-ZEUS only affects regions of strong compression and is therefore not crucially acting to damp hydrodynamical instabilities associated with the cloud.

Gas_10AV3, see Table 3.1 for viscosity values. A simulation using the Balsara switch but with the standard ($\alpha = 1.0, \beta = 2.0$) was also performed. Fig. 3.11 shows the outcome of the simulations at $t = 0.25, 0.75, 1.5$ and $2.25\,\tau_{\rm KH}$. We can directly see the impact these terms have on the stability of the simulation. The standard $\alpha = 1.0$, $\beta = 2.0$ is the most stable one, most probably due to the unphysical use of the α bulk viscosity. The use of $\alpha = 0$ and $\beta = 2.0$ or the Balsara switch renders very similar visual results. This is because the Balsara switch turns of viscosity where $|\nabla\cdot\boldsymbol{v}|/(|\nabla\cdot\boldsymbol{v}+|\nabla\times\boldsymbol{v}|)$ is significant, which is the case for shearing flows like on the surface of the cloud. Note that this is a very noisy quantity when measured using only 32 neighbours. By further lowering the shock capturing β viscosity we make the cloud even more unstable but it is not clear how physical this solution is. The shock front gets more blurred and we see strong post shock ringing effects. The reason for the increased instability in the $\alpha = 0$, $\beta = 0.5$, and $\alpha = 0$, $\beta = 0.1$ case is most probably due to high speed particles traveling through the poorly captured shock region and transferring momentum inside the cloud, perturbing it in an unphysical way.

We have performed simulations similar in spirit to the SPH ones using ENZO-ZEUS. There is formally no need for linear viscosity using this method except for hyper-sonic flows, but it is interesting to study the effect of lowering $Q_{\rm AV}$ in the same way as β. Fig 3.12 shows density slices from these simulations at $t = 1.5$ and $2.25\,\tau_{\rm KH}$. We see no impact on the cloud fragmentation except for minor morphological differences expected in turbulent regimes: $Q_{\rm AV}$ only serves to broaden the shock. Viscosity in grid based techniques are not as fundamental as in SPH techniques, where it must be set large enough to properly reproduce the behaviour of a fluid and not a collection of particles.

We chose not to experiment with linear viscosity in ENZO-ZEUS as it is truly a viscous term with the same functional form as what is associated with the SPH α viscosity *but* is also sensitive to expansion. A comparison can therefor not be made on equal terms.

To conclude, we see from these simulations how lowering the artificial viscosity outside shock regions will make the cloud in the SPH simulations less stable while losing the fluid behaviour for very low values. We still can't obtain agreement with the grid based codes. which leads us to suspect that there are more fundamental reasons behind the discrepancies.

Resolving Instabilities — In order to create an even simpler test problem to compare instabilities between codes, we carried out a classical Kelvin-Helmholtz test using GASOLINE and ENZO. We looked at the shearing

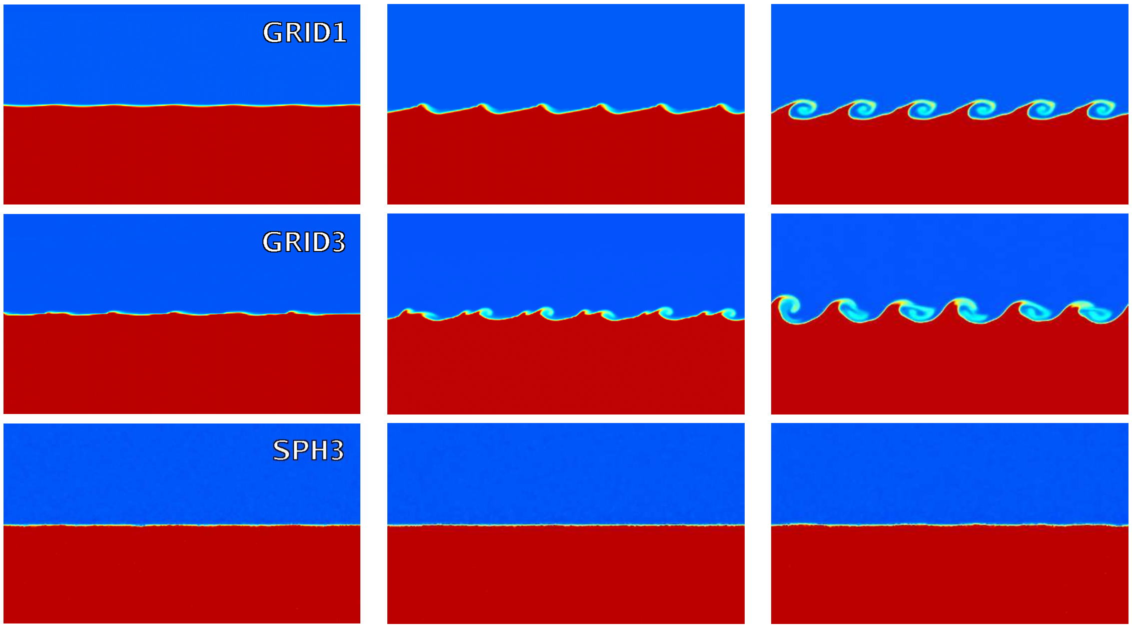

Figure 3.13: Density slices of, from top to bottom, GRID1, GRID3 and SPH3. The panels show the KH simulation at $t = \tau_{\rm KH}/3$, $2\tau_{\rm KH}/3$ and $\tau_{\rm KH}$. The grid simulations show clear growth of the KHI while this is completely absent in SPH.

Table 3.2: Performed KH runs

Resolution	χ	$\delta v/v_{\rm shear}$	IC	Name
ENZO				
$\{256, 256, 8\}$	8.0	$1/80$	lattice	GRID1
$\{256, 256, 8\}$	10.0	$1/40$	Poisson	GRID2
$\{256, 256, 8\}$	10.0	$1/40$	glass	GRID3
GASOLINE				
900 k part	8.0	$1/80$	lattice	SPH1
1.1 M part	10.0	$1/40$	Poisson	SPH2
1.1 M part	10.0	$1/40$	glass	SPH3

motion of two gases of different densities and with small perturbations imprinted at the boundary. This captures the hydrodynamics at the surface of the cloud in the blob test.

The setup is a periodic box with dimensions $\{L_x.L_y, L_z\} = \{1, 1, 1/32\}$, divided into two regions: one cold, high density and one warm, low density. The density and temperature ratio is $\chi = \rho_{\rm b}/\rho_{\rm t} = T_{\rm t}/T_{\rm b} = c_{\rm t}^2/c_{\rm b}^2$, putting the whole system in pressure equilibrium. The two layers are given constant and opposing shearing velocities, with the top layer moving leftward at a Mach number $\mathcal{M}_{\rm t} = v_{\rm t}/c_{\rm t} \approx 0.11$ and the bottom layer moving rightward at a Mach number $\mathcal{M}_{\rm b} = v_{\rm b}/c_{\rm b} \approx 0.34$ in the case of $\chi = 10$. The shear velocity becomes $v_{\rm shear} = 0.68\, c_{\rm b}$ and the subsonic regime will assure growth of instabilities (Vietri et al., 1997). This setup should mimic the growth of instabilities on the cloud surface.

To trigger instabilities we have imposed sinusoidal perturbation on the vertical velocity of the form

$$v_y(x) = \delta v_{\rm y} \sin(\lambda 2\pi x), \tag{3.19}$$

where δv_{y} is the amplitude of the perturbation in terms of the sound speed c_{b} and λ is the wavelength of the mode which we have put to $1/6$ in all of out tests. The perturbation is limited to a central strip around the interface of thickness 5% of the box size.

The initial conditions are again generated using particles for SPH. These are then mapped to a grid as explained in Section 3.4 to be used in the grid code, allowing a similar starting point for both codes. An important issue for this type of test is how the initial particles are distributed since this will introduce a certain amount of noise via discreteness. The most common techniques for this are:

- **Lattice:** Particles are ordered in a perfect grid. For a shearing layer test, this type of IC is optimal for grid codes as it traces the computational grid perfectly and suppresses all local density fluctuations.
- **Poisson:** Particles are randomly distributed to generate the IC of our problem. This type of setup generates local density variations, causing spurious pressure forces.
- **Glass:** A Poisson particle distribution, with our IC setup, is heated and relaxed until random velocities arising from pressure fluctuations are much smaller ($\lesssim$ few %) than the later imposed shear velocity (see Section 3.4).

Any initial condition with local density variations will trigger small scale KHIs. We carried out this test using all three methods in order to illustrate their impact. The lattice is obviously perfect for grid codes, making a perfectly homogeneous gas. This quality does not automatically produce clean SPH initial conditions due to the averaging over nearby particles. The Poisson ICs are very noisy in both the grid and SPH case, even though grid codes tend to smooth the noise over the cell sizes. The glass IC is intuitively the closest IC for both methods producing a self-consistent and homogeneous initial state for SPH simulations while leaving only small fluctuations for both grid and SPH methods.

This set of simulations and their characteristics are summarized in Table 3.2 and Fig. 3.13 shows the results, from top to bottom, GRID1, GRID3 and SPH3. We choose to show only one of the SPH results since all of these runs give the same result. GRID1 and GRID3 illustrate the difference between a highly idealized smooth setup (GRID1) and one with small scale noise (GRID3).

GRID1 nicely produces the KHIs and the growth time is in excellent agreement with that expected from Eq. 3.8. This growth is not as clean in GRID3, which is to be expected due to local noise in density which alters the visual outcome. However the KHI is still well resolved and the growth time is comparable to the analytical expectation.

The outcome of the SPH simulation is again very different from the grids. Perturbations are damped out very quickly both in velocity and density *regardless of choice* of initial conditions, resolution, perturbation strength and viscosity. We conclude that SPH in the form used in astrophysical simulations to-date is unable to capture dynamical instabilities such as KH when density gradients are present. As we will show in the next section, the reason for this stems from the way hydrodynamical forces are calculated in SPH in regions with strong gradients.

Mind the Gap — Fig. 3.14 shows a closeup of the SPH particles at the interface of the two fluids in SPH3 at $t = \tau_{\mathrm{KH}}$. There is a gap between them that has the size of an SPH smoothing kernel radius ($\sim 2h_{ij}$). This gap repeats periodically in each fluid, being smaller in the higher density fluid since the smoothing length (mean distance to the nearest 32 particles) is smaller there. This feature is found in *all* of our SPH KH simulations. It occurs very quickly and becomes more prominent with time. This phenomenon has been discussed before in the literature (e.g. Ritchie & Thomas, 2001; Tittley et al., 2001; Okamoto et al., 2003), especially in the context of numerical over-cooling (Pearce et al., 1999; Thacker et al., 2000; Springel & Hernquist, 2002; Marri & White, 2003) but no relation to resolving instabilities has been mentioned.

The gap can also be clearly seen in the cloud test simulation Fig. 3.5. Even though the gas is streaming with high velocity onto the leading surface of the cloud, spurious pressure forces prevent it from making any physical contact. The reason that the cloud loses mass in the SPH simulation is due to the vacuum behind the cloud into which the cloud expands from its edges. Here the gradients become smooth and the gas can be removed by the pressure difference between the cloud and the ambient medium that streams past.

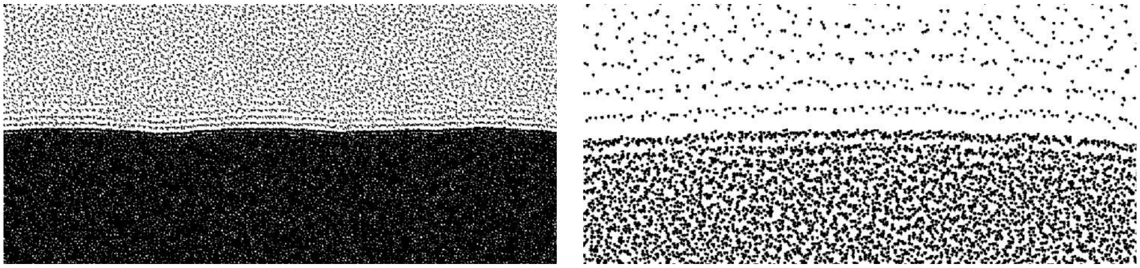

Figure 3.14: A close up view of the SPH particles at the boundaries between the shearing layers (left) and closer zoom in (right) for SPH3 at $\tau_{\rm KH}$. We can clearly see empty layers formed through erroneous pressure forces due to improper density calculations at density gradients. Even though the two fluids are moving relative to each other, the gap is so large that proper fluid interaction is severely decreased or even absent.

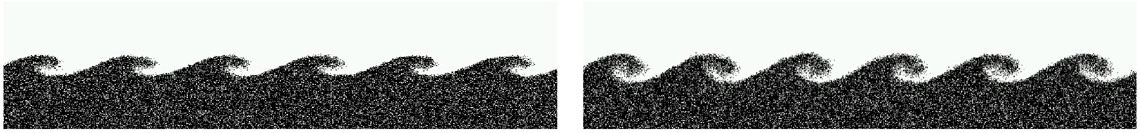

Figure 3.15: A zoom in of the SPH particles at the boundaries between the shearing layers for the isodensity SPH run with standard viscosity (left) and low viscosity (right) at $\tau_{\rm KH}$. The black and white region are particles that belonged to the initially separated shearing layers. We clearly see growth of the KHI in the standard implementation of SPH, and even stronger for the low viscosity version. The simulation was performed with GASOLINE using 10^6 particles in the same way as SPH3 described in Section 3.6.

The effect can be explained in the following way: Eq. 3.14 is the force on each SPH particle coming from the summation over the 32 nearest neighbours. The pressure is given by $P \sim \rho T$ in the assumed case of an ideal gas. This force calculation formally assumes that temperature, and more importantly, density gradients are small within the smoothing kernel, where temperature is a quantity accumulated over time while density usually is re-estimated at each time step. When a particle from a hot low density region approaches a cold high density region it will suddenly find a lot of neighbours at the edge of the smoothing sphere within the dense medium and its density will be overestimated. This leads to, through momentum conservation, a repulsive, fictitious, force on the particle, causing it to bounce back into the low density region. This behaviour leads to the formation of a gap between the two phases of the size $\sim 2h_{ij}$, where h_{ij} is the effective smoothing kernel length, either obtained by using smoothing length or smoothing kernel averaging (Hernquist & Katz, 1989), depending on the SPH implementation. Hot particles close to this gap will now have a strongly asymmetric distribution of particles around them resulting in an average pressure force pointing back into the vacuum layer. Particles then travel back into the empty region and the whole process is repeated. This particle migration and its associated pressure forces will act as an effective restoring force for the surface, a kind of tension. This together with the gap essentially removes multiphase behaviour from SPH. From the above arguments it is straightforward to see that in all standard formulations of SPH, any relaxed multi-phase particle distribution must have an associated gap.

As mentioned above, this erroneous treatment of density contrasts has also been found to produce overcooling in galaxy formation simulations. Tittley et al. (2001) showed that in subsonic regimes this behaviour leads to fictitious accretion of particles on the lateral sides of gas clouds such as the simulations showed in this paper. Solutions to this problem has been attempted by several authors (e.g. Ritchie & Thomas, 2001; Marri & White, 2003) by reformulating SPH to more accurately treat the particle interactions at steep boundaries. While this seems to remove the gap to some extent, it is unclear how this will affect the simulations discussed here. Possible solutions to the problem, such as improving the method of calculating gradients and minimizing their errors in SPH will be presented in a follow up paper by Read et al. (in preparation).

That erroneous density gradients are the root of the instability suppression becomes even more apparent by studying the KHI using a density contrast $\chi = 1$, in which the gap can not form. We performed a simulation using GASOLINE in the same way as SPH3 described in Section 3.6 but now using 10^6 particles $\chi = 1$. With this vanishing density gradient, SPH is able to capture the KHI, see Fig. 3.15. The left panel shows the KHI at $t = \tau_{\rm KH}$ for the standard $\alpha = 1.0$, $\beta = 2.0$ setting and the right panel shows the same time-step but using $\alpha = 0.01$ and $\beta = 1.0$. The less evolved standard viscosity simulation points out the effects of viscosity discussed in Section 3.6. Similar results have been recently found by Junk et al. (in preparation).

3.7. Summary

In this paper we have carried out hydrodynamical simulations of a cold gas cloud interacting with an ambient hot moving gas using state of the art simulations codes. Striking differences were found between the two main techniques for simulating fluids. While grid codes are able to resolve and treat dynamical instabilities and mixing, these processes are poorly or not at all resolved by the current SPH techniques. We show that the reason for this is that SPH, at least in the standard usage and formulation, inaccurately handles situations where density gradients are present. In these situations, SPH particles of low density close to high density regions suffer erroneous pressure forces due to the asymmetric density within the smoothing kernel. This causes a gap between regions of high density contrast, essentially decoupling the different phases of the fluid.

This behaviour has implications for many astrophysical situations. The stripping of gas from galaxies moving through a gaseous medium has already been discussed in the literature. The origin of disk galaxies is an important unsolved problem. Perhaps the inability to disrupt accreting gas clouds is one reason why numerical calculations have failed to produce pure disk systems. Simulating star formation regions and feedback processes also relies on the correct ability to model turbulence and interacting multiphase fluids.

It should be noted that the behaviour of the grid and SPH methods agree on timescales shorter than those of typical dynamical instabilities such as the Kelvin-Helmholtz and Rayleigh-Taylor instabilities. In our specific test of a cold cloud engulfed in a hot wind, there is good agreement in the early gas stripping phase occurring due to pressure differences arising in the Bernoulli zones. As soon as the large scale instabilities have grown, the results of the different methods diverge. There are several possible solutions to this behaviour in SPH calculations which we will explore in a separate work.

Acknowledgments

We acknowledge support from the European Science Foundation who funded an exploratory workshop in Wengen 2004 at which these tests were first discussed. FM and LM acknowledge support by the Swiss Institute of Technology through a Zwicky Prize Fellowship. OA would like to thank Alessandro Romeo, Peter Englmaier and Mordecai-Mark Mac Low for valuable discussions. AJG acknowledge the support from the Polish Ministry of Science through the grant 1P03D02626 and from the European Community's Human Potential Programme through the contract HPRN-CT-2002-00308, PLANETS. The AMR software (FLASH) used in this work was in part developed by the DOE-supported ASC / Alliance Centre for Astrophysical Thermonuclear Flashes at the University of Chicago. The FLASH calculations were performed at the Interdisciplinary Centre for Mathematical and Computational Modeling in Warsaw, Poland.

The ENZO, GASOLINE and GADGET-2 simulations, as well all of the analysis presented in this paper, were performed on the zBox1 and zBox2 supercomputers (http://krone.physik.unizh.ch/∼dpotter/zbox/) at the University of Zürich.

4.

Large Scale Galactic Turbulence

Can Self-Gravity Drive the Observed HI Velocity Dispersions?[25]

> ‘When I meet God, I am going to ask him two questions: Why relativity?
> And why turbulence? I truly believe he will have an answer for the first.’
>
> — *Werner Heisenberg, 1976*

Observations of turbulent velocity dispersions in the HI component of galactic discs show a characteristic floor in galaxies with low star formation rates and within individual galaxies the dispersion profiles decline with radius. We carry out several high resolution adaptive mesh simulations of gaseous discs embedded within dark matter haloes to explore the roles of cooling, star-formation, feedback, shearing motions and baryon fraction in driving turbulent motions. In all simulations the disc slowly cools until gravitational and thermal instabilities give rise to a multiphase medium in which a large population of dense self-gravitating cold clouds are embedded within a warm gaseous phase that forms through shock heating. The diffuse gas is highly turbulent and is an outcome of large scale driving of global non-axisymmetric modes as well as cloud-cloud tidal interactions and merging. At low star-formation rates these processes alone can explain the observed HI velocity dispersion profiles and the characteristic value of $\sim 10\,\mathrm{km\,s^{-1}}$ observed within a wide range of disc galaxies. Supernovae feedback creates a significant hot gaseous phase and is an important driver of turbulence in galaxies with a star-formation rate per unit area $\gtrsim 10^{-3}\,M_\odot\,\mathrm{yr^{-1}\,kpc^{-2}}$.

4.1. Introduction

The interstellar medium (ISM) is dominated by irregular/turbulent gas motions (e.g. Larson, 1981; Elmegreen & Scalo, 2004). HI emission lines in most spiral galaxies have characteristic velocity dispersions of $\sigma \sim 10\,\mathrm{km/s}$ on a scale of a few hundred parsecs, exceeding the values expected from purely thermal effects. The data in Fig. 4.1, assembled by Dib et al. (2006), also shows a transition to much larger values in active/starbursting galaxies. Recent high resolution observations by Petric & Rupen (2007) of the nearly face on disc galaxy NGC 1058 (see also Dickey et al., 1990) provides us data on the radial behavior of the vertical velocity dispersion. They find that the dispersion declines with radius from $\sim 12 - 15\ \mathrm{km\,s^{-1}}$ in the inner parts to $\sim 4 - 6\ \mathrm{km\,s^{-1}}$ in the outer and is uncorrelated with active regions such as star formation sites and spiral arms. This is attributed to small scale ($< 0.7\,\mathrm{kpc}$) bulk motions. Petric & Rupen state that any model attempting to explain turbulence in the ISM *must* also explain the radial decline that also has been detected in previous studies of e.g. NGC 6946 (Boulanger & Viallefond, 1992), NGC 628 (Kamphuis & Sancisi, 1993; van der Hulst, 1996), NGC 2915 (Meurer et al., 1996).

The main source(s) of energy driving the ISM dynamics is still not clear (Burkert, 2006), even though there are several candidates capable of driving the ISM turbulence (Mac Low & Klessen, 2004). A commonly discussed source is of stellar origin i.e. large-scale expanding outflows from high-pressure HII regions (Kessel-Deynet & Burkert, 2003), stellar winds or supernovae. Whilst supernovae explosions might dominate the energy input into the ISM (e.g. Mac Low & Klessen, 2004; Dib et al., 2006), the mechanism is unable to

---

[25] This chapter has been published as Agertz et al. (2009) in Monthly Notices of the Royal Astronomical Society, Volume 392, Issue 1, pp. 294-308.

explain the broad HI lines in galaxies with a low star formation rate (SFR) and in regions of moderate stellar activity as in the outer parts of disc galaxies. Many numerical studies have been carried out to understand the influence of supernovae in galactic discs (e.g. Kim et al., 2001; de Avillez & Breitschwerdt, 2004, 2005; Slyz et al., 2005; Mac Low et al., 2005; Joung & Mac Low, 2006). Dib et al. (2006) reproduced the starbursting transition seen in Fig. 4.1 but was unable to produce velocity dispersions larger than $\sim 3\,\mathrm{km\,s^{-1}}$ for low values of SFR/Area. This strongly suggests that something else is contributing to the energy budget. In addition, large scale holes, usually attributed to correlated supernovae explosions (e.g. Puche et al., 1992), are in some cases surprisingly *uncorrelated* to stellar activity (Rhode et al., 1999).

Another source of turbulence is galactic rotation. This is a huge reservoir of energy (Fleck, 1981) and any mechanism able to generate random motions from ordered circular motion could sustain turbulence for many orbital times. Numerical work of Wada et al. (2002) and Wada & Norman (2007) has shown that realistic global models of galactic discs form a very complicated turbulent velocity field associated with a multiphase ISM. The only active source for this is shear coupled to gravitational and thermal instability. Local isothermal simulations of the ISM done by Kim & Ostriker (2007) (also previous work e.g. Kim & Ostriker (2001) and Kim et al. (2003)) support this notion. They demonstrated that gas in a marginally stable galactic discs obtains, under certain conditions, velocity dispersions as large as the sound speed (here $c_s = 7\,\mathrm{km\,s^{-1}}$) due the swing-amplifier (Goldreich & Lynden-Bell, 1965b; Julian & Toomre, 1966; Toomre, 1981; Fuchs, 2001). The swing-amplifier is when a leading wave is amplified into a trailing wave. The underlying mechanism is shear and self-gravity.

Fukunaga & Tosa (1989) showed that rotational energy randomizes the motions of the cold cloud component of a galactic disc via gravitational scattering from their random epicyclic motions. This was was later quantified by Gammie et al. (1991) who showed that the cloud velocity dispersion could reach $\sim 5-6\,\mathrm{km\,s^{-1}}$ in this way, in agreement with observations (Stark & Brand, 1989). We will discuss this mechanism and its impact on the ISM in more detail in Sect.4.3.

The Magneto-Rotational-Instability (MRI) (Balbus & Hawley, 1991; Sellwood & Balbus, 1999) coupled with galactic shear is also a possible driver of turbulence. Piontek & Ostriker (2004) and Piontek & Ostriker (2005) obtained reasonable values of $\sim 8\,\mathrm{km\,s^{-1}}$ under favorable conditions. This mechanism becomes significant at low densities and might be important in the more diffuse outer part of galaxies.

In this paper we carry out high-resolution 3-dimensional Adaptive Mesh Refinement (AMR) simulations to form a realistic multiphase ISM in which we can disentangle the contributing effects of self-gravity and supernovae driven turbulence. The simulations incorporate realistic prescriptions for cooling, star formation and supernovae feedback. Similar numerical simulations have been carried out before (e.g. Gerritsen & Icke, 1997; Wada et al., 2002; Bottema, 2003; Tasker & Bryan, 2006; Wada & Norman, 2007) but without addressing directly the issues discussed in this paper.

The paper is organized as follows. In Sect. 4.2 we describe the numerical method used for this work and the setup of the galactic discs. In Sect. 4.3 we present the results from the numerical simulations, where the results treating the turbulent ISM are given in Sect.4.3. Sect. 4.4 summarizes and discusses our conclusions.

4.2. Numerical Modeling

The Code and Subgrid Modeling — We use the adaptive mesh refinement (AMR) hydrodynamics code RAMSES (Teyssier, 2002). The code uses a second order Godunov scheme to solve the Euler equations. The equation of state of the gas is that of a perfect mono-atomic gas with an adiabatic index $\gamma = 5/3$. Self-gravity of the gas is calculated by solving the Poisson equation using the multigrid method (Brandt, 1977) on the coarse grid and by the conjugate gradient method on finer ones. The collisionless star particles are evolved using the particle-mesh technique. The dark matter is treated as a smooth background density field that is added as a static source term in the Poisson solver. The code adopts the cooling function of Sutherland & Dopita (1993) for cooling at temperatures $10^4 - 10^{8.5}\,\mathrm{K}$. We extend cooling down to 300 K using the parametrization of Rosen & Bregman (1995). The effect of metallicity is approximated by using a linear scaling of the functions.

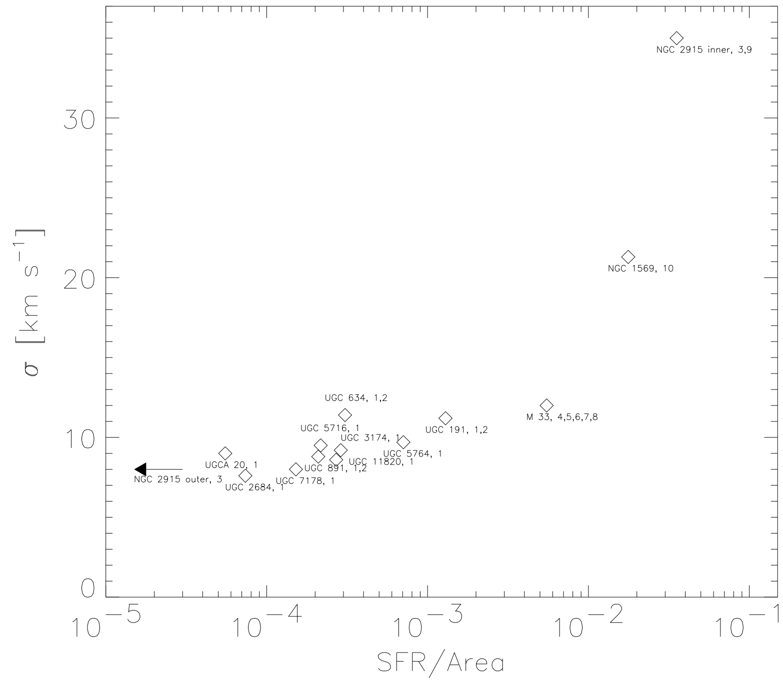

Figure 4.1: Characteristic HI gas velocity dispersion of a sample of galaxies as a function of the derived star formation rate in units $M_{\odot}\,\mathrm{yr}^{-1}\,\mathrm{kpc}^{-2}$ as plotted by Dib et al. (2006). The figure is reproduced here by courtesy of Sami Dib, Eric Bell and Andreas Burkert, and by permission of the AAS.

The star formation recipe is described in Dubois & Teyssier (2008) but we summarize the main points here for completeness. In a cell, gas is converted to a star particle using a Schmidt law

$$\dot{\rho}_* = -\frac{\rho}{t_*} \text{ if } \rho > \rho_0 \qquad \dot{\rho}_* = 0 \text{ otherwise}, \tag{4.1}$$

where t_* is the star formation time scale and ρ_0 is an arbitrary threshold that should be chosen to carefully make physical sense when related to the resolution and cooling floor. The star formation timescale is related to the local free-fall time,

$$t_* = t_0 \left(\frac{\rho}{\rho_0}\right)^{-1/2}. \tag{4.2}$$

The parameters ρ_0 and t_0 are in reality scale dependent and not very well understood theoretically. A common way to get around this is to calibrate them to star formation rates in local galaxies i.e. to the Kennicutt (1998) law and make sure that the values are compatible with modern estimates of star formation efficiencies (Krumholz & Tan, 2007) of $\sim 1-2\,\%$ per free-fall time in giant molecular clouds (GMCs). For example, if the star formation threshold $\rho_0 = 100\,\mathrm{cm}^{-3}$, the free fall time is $5\,$Myr meaning we can use $t_0 = 250\,$Myr to get $2\,\%$ efficiency per free fall time. As soon as a cell is eligible for star formation, particles are spawned using a Poisson process where the stellar mass is connected to the chosen threshold and code resolution (see Dubois & Teyssier, 2008).

The implementation of supernovae feedback is also described in the above reference (see their Appendix A). In the simulations that include feedback we assume that $50\,\%$ of the total supernovae energy, $E_{SN} = 10^{51}$ ergs, goes into thermal energy where $\eta_{\mathrm{SN}} = 10\,\%$ of each solar mass of stars that is formed is recycled as supernovae ejecta. The energy and gas release is also delayed by $10\,$Myr from the time of explosion by creating debris particles on at the time of explosion. By delaying the energy and mass release we allow for it to take place outside of dense environments, hence preventing it from radiating away too quickly. We follow

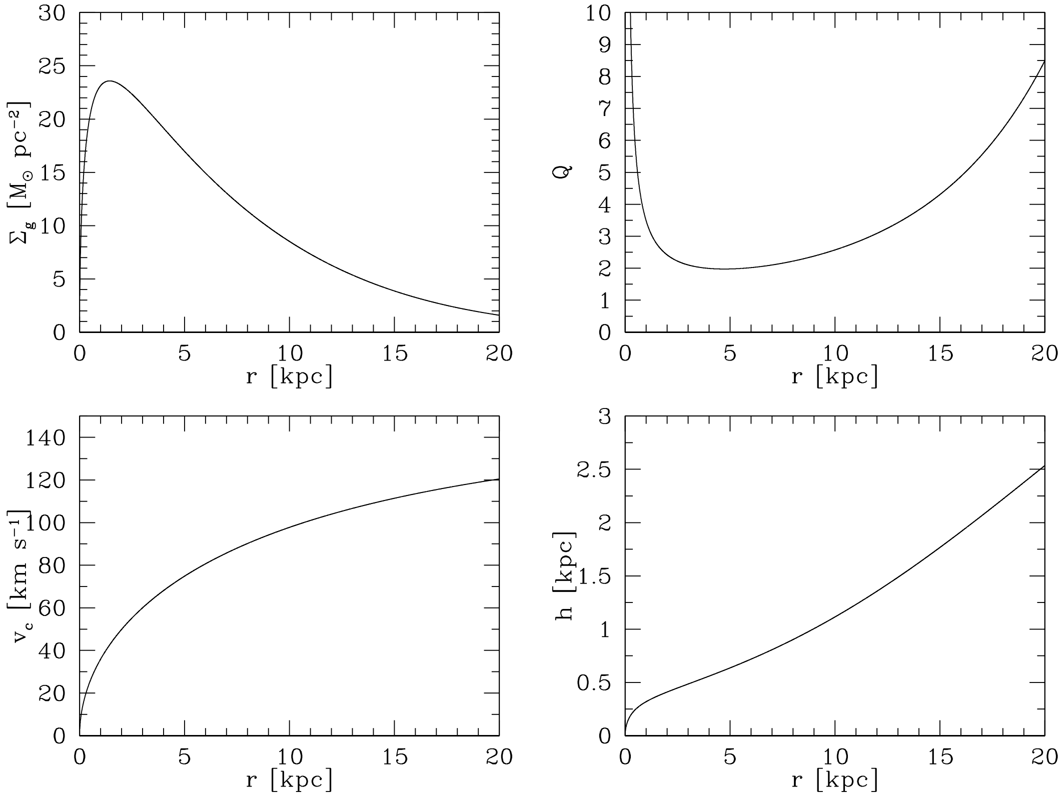

Figure 4.2: Characteristics of the initial gas distribution in all simulations, apart from RUN6 and RUN8. The panels show, from top left to bottom right, the radial dependence of the surface density, the gaseous Q-value, the circular velocity and the scale height.

the prescription of Dubois & Teyssier (2008) and apply a large mass loading factor for the debris particles, i.e. $\eta_W = 1.0$. We use thermal rather than kinetic energy releases since we resolve the clumpy ISM and follow shocks self-consistently. In addition, a model that allows for debris particles to transfer kinetic energy is no longer valid as a Sedov explosion assumes a homogeneous medium to propagate into. While any treatment is inherently sub-grid, the supernovae impact should converge with enough resolution (Ceverino & Klypin, 2007).

In order to model subgrid gaseous equation of states and to avoid artificial gas fragmentation, the gas is given a polytropic equation of state as it crosses ρ_0. The temperature is set to

$$T = T_0 \left(\frac{\rho}{\rho_0}\right)^{\gamma_0 - 1}. \tag{4.3}$$

T_0 is set to be the cooling floor of our simulations for consistency and $\gamma_0 = 2.0$.

Initial Conditions — Our initial condition (IC) is an axisymmetric galactic gas disc in equilibrium with an NFW (Navarro et al., 1997) dark matter halo. All relevant IC characteristics are presented in Fig. (4.2). The disc is initially isothermal at $T = 10^4$ K having an exponential density profile, in cylindrical coordinates r and z,

$$\rho(r, z) = \mathrm{sech}^2(z/h(r))\rho_0 e^{-r/r_0}, \tag{4.4}$$

Table 4.1: Performed simulations.

Simulation	Min. Δx	ρ_0 [cm^{-3}]	Cooling floor, T_0	Modelling comments
RUN0	24 pc	–	300 K	Hydrodynamics with self-gravity
RUN1	24 pc	10	300 K	Like RUN0 + star formation
RUN2	6 pc	100	300 K	Like RUN1
RUN3	24 pc	10	300 K	Like RUN1 + supernova feedback
RUN4	24 pc	–	1 000 K	Like RUN0
RUN5	24 pc	–	10 000 K	Like RUN0
RUN6	24 pc	10	300 K	Like RUN1 but with 1/3 of gas mass
RUN7	24 pc	–	5 000 K	Like RUN0
RUN8	24 pc	10	300 K	Like RUN1 but higher halo concentration

where r_0 is the scale radius and $h(r)$ the scale height. The sech^2-term owes to the isothermality of the gas such that

$$h(r) = \frac{c_s^2}{\pi G \Sigma(r)} \tag{4.5}$$

where c_s is the local sound speed and $\Sigma(r)$ the local total surface density ($\Sigma = \Sigma_{\rm gas} + \Sigma_{\rm DM}$), naturally leading to a flaring disc. Experiments using a radial $1/r$ gas distribution were also performed (not discussed further) without any significant differences.

We choose to model an M33 type galactic disc as it is a nearby well observed gas rich system. All global characteristics of the initial disc are in agreement with the observations presented by Corbelli (2003). M33 has a total gas mass (HI + HII + He) of $\sim 3.2 \times 10^9 M_\odot$ and an estimated stellar mass of $3 - 6 \times 10^9 M_\odot$. We choose the initial gas mass to be in the high end of the total baryonic mass, i.e. $\approx 9.2 \times 10^9 M_\odot$ as a lot of the outer material will not be a part of what we would associate with a galaxy and is only an artifact of the way we model isolated discs. The gas is assumed to have a mean metallicity of $0.3\,Z_\odot$.

We initialize the disc in a stable configuration where most of the disc has a Toomre parameter $Q \sim 2 - 3$. These large values of Q are desirable as we want the cooling to initiate instabilities and not our choice of initial conditions. A fairly large scale radius of $r = 4.0\,\mathrm{kpc}$ is used. As this only reflects the very early setup of a forming disc galaxy this will not be an issue in our modeling. The dark matter halo has a concentration of $c = 8.0$, scale radius $R_s = 35.0\,\mathrm{kpc}$ and a total mass of $10^{12} M_\odot$. These model parameters can be perceived as odd but is necessary for a best fit NFW-halo which, in accordance with observations (Corbelli, 2003), reproduce a dark halo mass that within 17 kpc is $\sim 5 \times 10^{10}\,M_\odot$. This mass sets a lower limit on the actual dark matter halo mass. The HI velocity profile is still (slowly) rising at this radius. Scenario with different mass profiles, gas masses, shear and cooling floors are also explored.

Numerically this setup is initialized at a resolution of $100\,\mathrm{pc}$ using a nested hierarchy of grids situated in a simulations cube of size $L_{\rm box} = 200\,\mathrm{kpc}$. We achieve higher resolution by refining cells both on based on a density and Jeans mass criterion (see Sect. 4.2 for details). The maximum allowed resolution is indicated in Table 1.

Simulation Suite — The performed runs are listed in Table 1. RUN0 serves as our base run where we only consider the self-gravitating cooling gas and dark matter. RUN1 introduces star formation as does RUN2 but at a higher resolution. RUN3 is identical to RUN1 but implements the feedback prescription described in Sect. 4.2. These are our four fiducial simulations to understand the importance of these physical mechanisms. To explore how choices of the gas cooling changes the outcome, RUN4, 5 and 7 are identical to RUN0 except for a truncation in the cooling function at the indicated thresholds. This will determine the ability of a disc to develop a gravitoturbulent state (see e.g. Gammie, 2001). RUN6 adopts 1/3 of the gas mass, making it a much more stable system. Finally, RUN8 adopts a very concentrated halo ($c = 40$, $R_s = 7\,\mathrm{kpc}$, $M = 3 \times 10^{11} M_\odot$) peaking at $120\,\mathrm{km\,s^{-1}}$ to assess the influence of a different shear, $d\Omega/dR$.

Numerical Considerations — In these types of experiments it is important to consistently resolve the Jeans scale associated with the chosen cooling floor. Truelove et al. (1997) demonstrated, using isothermal simulations, that at least 4 resolution elements are necessary to avoid artificial gas fragmentation. Our strategy is to choose realistic star formation density thresholds that together with the cooling temperature floor gives us a Jeans scale that can be resolved according to the Truelove criteria. The Jeans length given by

$$\lambda_{\mathrm{J}} = \sqrt{\frac{\pi c_{\mathrm{s}}^2}{G\rho}}. \tag{4.6}$$

In our simulations, where the temperature floor is set at $T = 300\,\mathrm{K}$, this can be rewritten as

$$\lambda_{\mathrm{J}} \approx 312\sqrt{1/n}\,\mathrm{pc} \tag{4.7}$$

where n is expressed in cm^{-3}. In order to satisfy the Truelove criterion we adaptively refine on a Jeans mass down to a resolution of $\Delta x = 24(6)\,\mathrm{pc}$ in RUN1 (RUN2) where we have set the star formation density threshold to $\rho_0 = 10(100)\,\mathrm{cm}^{-3}$. In addition to this precaution, the background ISM polytropic EOS (see Sect. 4.2) is activated at the same threshold, ensuring us that the Jeans scale never falls below the minimum value of $\sim 100(25)\,\mathrm{pc}$ set by Eq. 4.7 (at $n = 10(100)\,\mathrm{cm}^{-3}$). Additional simulations have been conducted adopting 16 cells per Jeans length to asses the fidelity of the gravitational fragmentation without any significant difference in outcome. No physical perturbations for gravitational instability are seeded in the initially smooth disc meaning the actual perturbations existing arise from the AMR grid. As we are only interested in the long-time ($t > 1\,\mathrm{Gyr}$) dynamical evolution of the system, the actual morphology of the early unstable disc is of little importance.

4.3. Results

As we will describe in Sect. 4.3, the response to gravitational instabilities, both in the form of bound structures and local non-axissymmetric instabilities, is an important source of turbulence. Therefore, before addressing the issue of turbulence we characterize the global gas evolution, phase-structure (density and temperature) and stability of the simulated galactic discs in Sect. 4.3 and Sect. 4.3.

General Evolution and Morphology —

☐ *Gas Evolution* — Fig. 4.3 shows the total gas surface density maps of RUN1, RUN2 and RUN3 at $t = 0.5, 1.0, 1.5$ and 2.0 Gyr. All simulations evolve in a similar fashion: the initial gas distribution cools down slowly, loses pressure support and contracts in the vertical direction. After a few 100 Myr the central part of the disc is cold enough to become gravitational and thermally unstable and fragments into bound clouds. This process quickly proceeds to larger radii. Non-axissymmetric instabilities such as swing amplification aids the process everywhere, especially in the outer parts where the gas is only mildly unstable. The formation of bound clouds and elongated structures such as shearing filaments is very similar to that found by e.g. Kim et al. (2003) and Kim & Ostriker (2007) for an unstable or marginally stable ISM. This clumpy structure is also visible in the evolution of the total gas surface density and rotational velocity in Fig. 4.4. The decrease of mean surface density is due to star formation. Fig. 4.4 also shows the contribution to the rotational velocity at $t = 1.0\,\mathrm{Gyr}$ from gas, stars and dark matter. We can clearly see that the initially gas and dark matter only system has evolved to a state in which the relative contributions and their magnitudes agree well with M33 observations (e.g Corbelli, 2003). The evolution of the total gas and stellar mass in RUN1 and RUN3 are shown in Fig. 4.5. We note that a long-time evolution ($t > 1.5\,\mathrm{Gyr}$) of the galactic discs will force them to move away from a gas rich system such as M33, having $\sim 30\,\% - 50\,\%$ of its baryonic mass in gas, approaching $\sim 20\,\% - 22\,\%$ at $t = 2\,\mathrm{Gyr}$. Taking gas infall into account and using a more realistic star formation prescription could remedy this.

We observe significant cloud-cloud and cloud-ISM interactions as the disc evolves. The clouds undergo both collisions leading to coalescence as well as tidal and long range interactions inducing torques into the gas. Shearing wavelets form out of the disc in between the cold clouds. These structures interact with each other as well as the clouds for the entire simulation period. The clumpy ISM acts as an effective viscosity (Lin & Pringle, 1987) forcing material to sink to the center whilst smaller clumps stay on more regular orbits at larger radii. This effect is stronger at early times when the typical cloud collision times-scale is short. In between

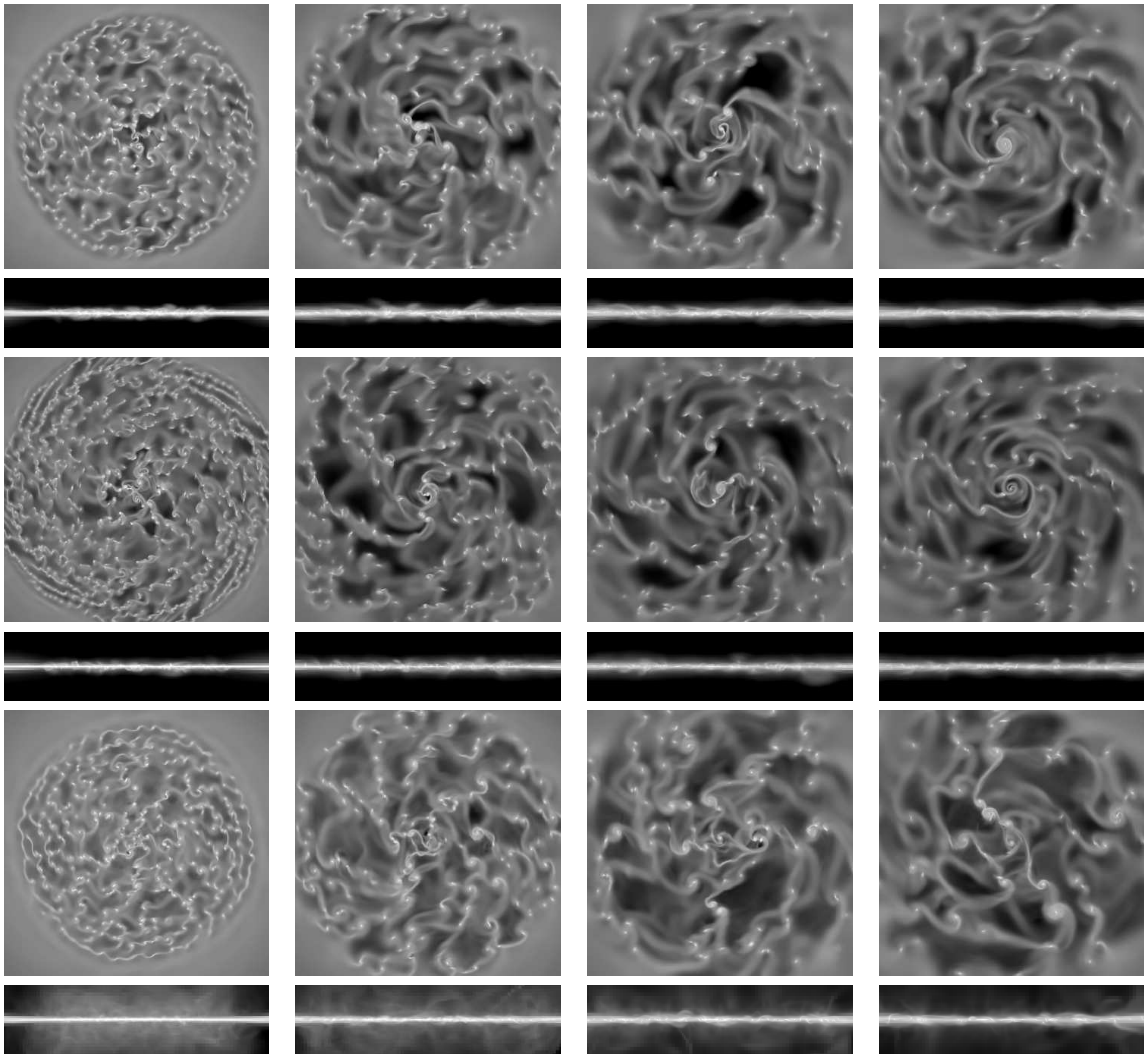

Figure 4.3: Logarithmic column density plots of the gas in the range $\Sigma_{\rm g} = 10^{18} - 10^{23}\,{\rm cm}^{-3}$. Each panel shows a face-on $30 \times 30\ {\rm kpc}^2$ map centered on the disc. The associated edge-on map is 8 kpc in height. From top to bottom we see RUN1, RUN2 and RUN3 at times, from left to right, $t = 0.5, 1.0, 1.5$ and $2.0\,{\rm Gyr}$.

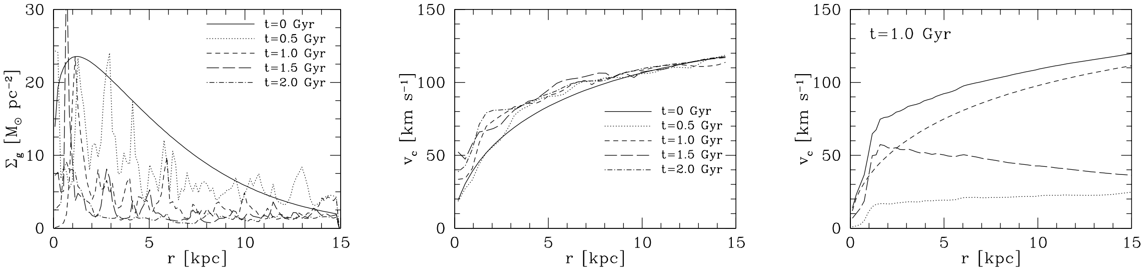

Figure 4.4: Time evolution of the surface density (*left*) and rotational velocity (*middle*) for the gas component in RUN1. The contributions to the rotational velocity (solid line) at $t = 1.0\,{\rm Gyr}$ (*right*) from the gaseous (dotted line), stellar (long-dashed line) and dark matter (dash line) components are in good agreement with observations of M33.

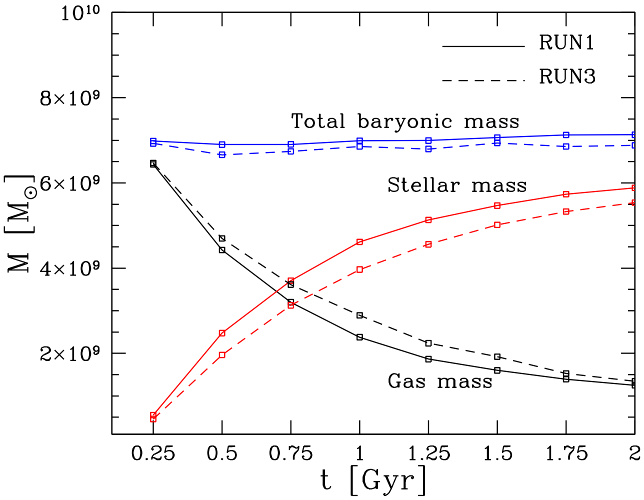

Figure 4.5: Time evolution of the mass of the gas and stellar component in a $30\,\mathrm{kpc}$ cube centered on the disc, as seen in Fig 4.3.

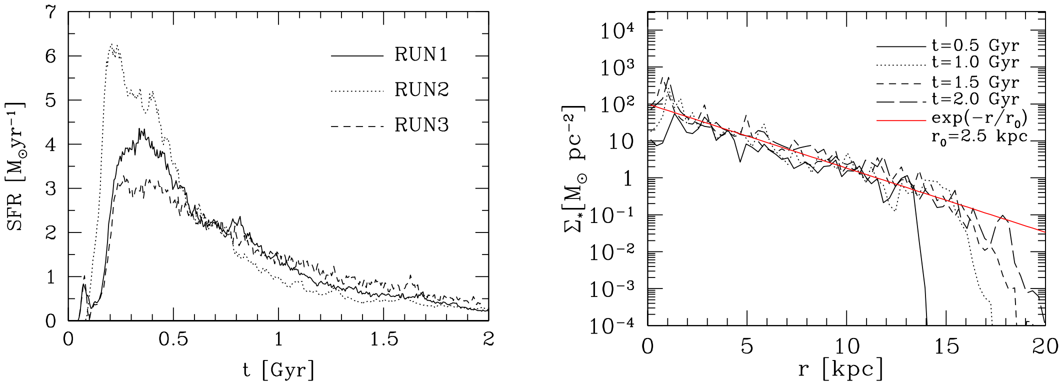

Figure 4.6: (*Left*) The star formation rate over time for RUN1, RUN2 and RUN3. The higher resolution in RUN2 allows for more star formation in the outer disc, while the feedback of RUN3 lowers the efficiency. At late times the SFRs show little difference. (*Right*) Evolution of the stellar surface density. The density profiles are at all times well fitted with an exponential function. The red line is for $r_0 = 2.5\,\mathrm{kpc}$.

dense clouds and stretched filaments, the ISM also develops under-dense regions ($\Sigma \lesssim 10^{18}\mathrm{cm}^{-2}$) on scales of $500\,\mathrm{pc}$ to several kpc. At later times, signatures of large scale spiral structure appear in the gaseous disc in which the cold clouds align.

In what way do the simulations differ? RUN1 and RUN2 evolve in a very similar fashion. However, the higher resolution in RUN2 means that further swing amplified instabilities can occur in the outer parts (see $t = 0.5$ Gyr in Fig. 4.3). Apart from this, the overall morphology and statistics are in good agreement throughout the whole simulation time, indicating convergence. The feedback in RUN3 successfully ejects hot low density gas into the ISM as well as out of the disc plane. This process is very efficient at early times when the SFR/Area is high but calms down as the SFR self-regulates, see Sect. 4.3 for discussion. The feedback also alters the structure of the disc. As seen in Fig. 4.3, the late time spiral patterns are not as pronounced as in RUN1 and RUN2 and fewer low-mass clouds have survived.

☐ *Star Formation* — The left panel in Fig. 4.6 shows the SFR of RUN1, RUN2 and RUN3. The main difference between the simulations is found in the most active star forming time, $t \sim 0.2 - 0.5\,\mathrm{Gyr}$, when

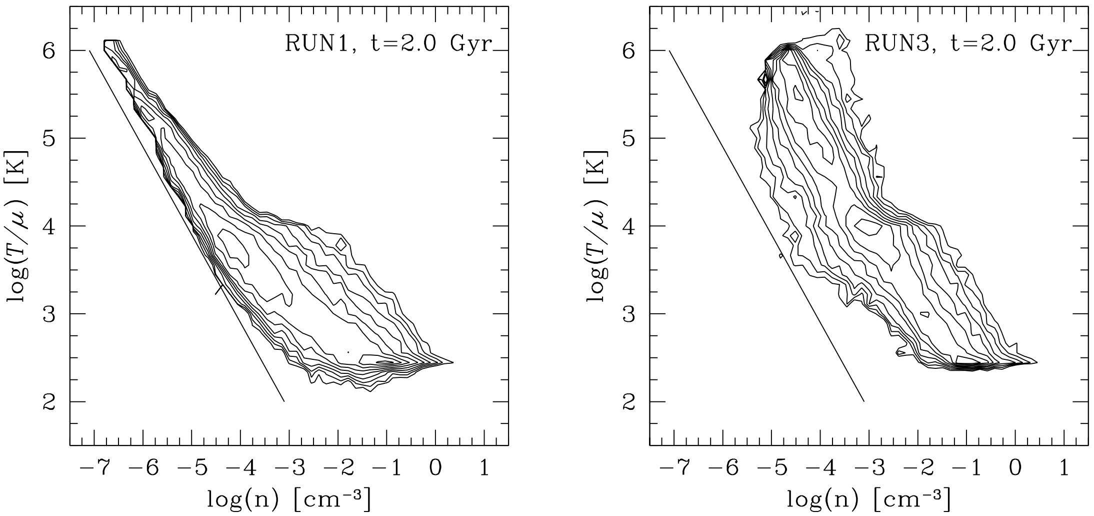

Figure 4.7: Phase diagrams for RUN1 (*left*) and RUN3 (*right*) at $t = 2.0\,\mathrm{Gyr}$. The solid straight line indicate the isobar.

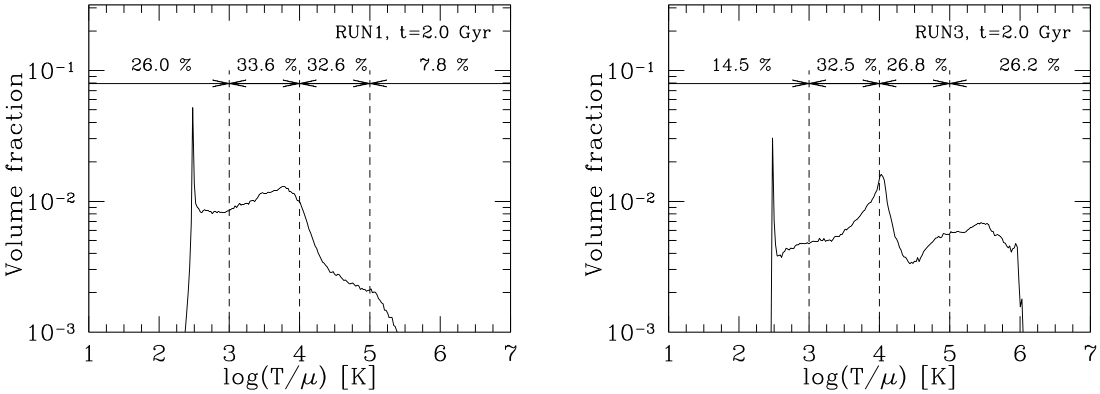

Figure 4.8: Volume fraction for RUN1 (*left*) and RUN3 (*right*) at $t = 2.0\,\mathrm{Gyr}$. While both models clearly show a cold and warm gas phase, the hot phase is only present in RUN3.

the initial gravitational and thermal instabilities have formed dense clouds. The higher resolution in RUN2 allows for cloud formation in the less dense outer parts of the disc, leading to a higher SFR. The SNe feedback in RUN3 dampens star formation as explosions heat or disperse star forming clouds. However, after this star-bursting period the disc settles to a quiescent phase where all simulation approach a SFR $\sim 0.25 M_{\odot}\mathrm{yr}^{-1}$.

The stellar surface density is at all times well-fitted with an exponential function, $\Sigma_*(r) \sim \exp(-r/r_0)$ with $r_0 \sim 2.5\,\mathrm{kpc}$, out to a truncation radius which grows with time, see right panel in Fig. 4.6. This may not come as a surprise as the initial gaseous disc is set up to be an exponential. At the end of the simulation time there are $\sim 1.3 \times 10^6$ stars in RUN1, 8.8×10^6 in RUN2 and 2.5×10^6 in RUN3.

Composition and State —

□ *A Multiphase ISM* — Theoretical models of the ISM (e.g. McKee & Ostriker, 1977b) have a three phase structure consisting of a cold, warm and hot phase in pressure equilibrium where regulation is obtained through the balance of radiative cooling and supernovae heating. More updated models separate the phases further based on their ionization state. In this work we refer to the phases as cold ($T \lesssim 10^3\,\mathrm{K}$), warm ($10^3\,\mathrm{K} \lesssim T \lesssim 10^5\,\mathrm{K}$) and hot ($10^5\,\mathrm{K} \gtrsim T$). A realistic ISM is very complicated which can be seen in the volume-weighted phase diagrams of RUN1 and RUN3 in Fig. 4.7. Both runs show a wide range of temperatures and densities. RUN3 clearly displays distinct cold, warm and hot phases aligning to an isobaric strip, i.e. $P \sim \rho T \sim \mathrm{constant}$. Around this region we observe a large spread in both temperature and density. This analysis is approximately valid for RUN1 even though the warm and hot gas are smeared over less distinct phase regions and sits at slightly lower densities compared to RUN3. The cold phase is almost identical in the two models. The existence of a hot tenuous phase in RUN3 is due to supernovae heating, but why do we find hot gas in RUN1? As the initial circular velocity is set for the whole computational domain, the low density ambient gas at $T = 10^4\,\mathrm{K}$ experiences a mild shock heating and settles into pressure equilibrium with the denser galactic disc. Some of the hot gas can also be found in the cloud-induced shocks in the disc.

Fig. 4.8 shows the volume fraction occupied by gas at different temperatures. We also indicate the total percentage of gas in different temperature regions. The expected two-phase structure in RUN1 is evident. We find a clearly peaked cold phase with a transition into a warm phase in between $1\,000$ K and $10\,000$ K, peaking at $6\,000 - 7\,000$K which is the thermally unstable regime. The origin of the warm phase is shock-heating. RUN3 shows the same cold phase but the warm phase now strongly peaks at $10\,000$ K, just at the maximum peak of the cooling function. A hot gas phase is clearly present even though very little gas exists above 10^6 K. As in RUN1, the warm phase dominates the gas volume.

It is desirable to approximately reproduce a mass distribution of molecular, cold atomic, warm atomic, warm ionized and hot ionized gas that agrees with observations (see e.g. Ferrière (2001) for the Galactic inventory). However, even among the local group spirals there can be significant differences between the phase-distributions. For example, M31 has $\sim 40\,\%$ of its gas in cold HI, Milky Way $\sim 25\,\%$ and M33 only $\sim 15\,\%$ (Dickey & Brinks, 1993). These differences could be due to the variation in baryon to dark matter fraction as we move down the Hubble sequence. As we show later, the formation of cold clouds is particularly sensitive to the gaseous disc mass. However, it is still instructive to compare our phase values of RUN1 and RUN3 at $t = 1.5\,\mathrm{Gyr}$ to those of Ferrière (2001) for the Milky Way. Roughly $50\,\%$ of the Milky Way gas is in molecular and cold atomic ($50 - 100\,\mathrm{K}$) clouds. Our simulations only allow for cooling down to 300 K and can hence not discriminate between the coldest gas phases. By labeling all dense gas of $T < 350\,\mathrm{K}$ as a joint cold cloud phase we find that $\sim 55(47)\,\%$ of the total gas mass in RUN1(3) is cold. The warm neutral gas phase ($10^3\,\mathrm{K} < \mathrm{T} < 10^4\,\mathrm{K}$) has $\sim 11(16)\,\%$ while the total neutral mass fraction of the gas outside of clouds ($350\,\mathrm{K} < \mathrm{T} < 10^4\,\mathrm{K}$) is $39(45)\,\%$ respectively. The former value is lower than the Milky Way value ($\sim 40\,\%$) which could be due to the fact that our initial conditions are more suitable for comparison with Sc galaxies. Also, including a homogenous UV background field in the simulations would heat the diffuse HI gas which seems to be the case in similar studies (e.g. Bottema, 2003). Furthermore, the observed mass of warm phase in the Galaxy is derived from the observed HI velocity dispersion of $6 - 9\ \mathrm{km\,s^{-1}}$ under the assumption of *only* thermal broadening (Ferrière, 2001). A turbulent component can allow for the existence of colder gas yet retaining the velocity dispersion values. We will explore this notion further in Sect. 4.3.

□ *Disc Stability* — To understand the relevance of gravitational instabilities in the simulated multi-phase discs, we use the Toomre parameter (Toomre, 1964) defined, for gas (Goldreich & Lynden-Bell, 1965a), as

$$Q_\mathrm{g} = \frac{\kappa c_\mathrm{s}}{\pi G \Sigma_\mathrm{g}}, \tag{4.8}$$

where c_s is the sound speed of the gas. Since the gas has turbulent motions it is appropriate to use the effective dispersion $\sigma_\mathrm{eff}^2 = c_\mathrm{s}^2 + \sigma_\mathrm{1D}^2$, where σ_1D^2 is the average of the full three-dimensional velocity dispersion. The Toomre parameter is valid for local axisymmetric perturbations of two-dimensional discs, where $Q_\mathrm{g} < Q_\mathrm{c} = 1$ implies instability. However, Q_g has been shown to characterize the response of discs to general gravitational instabilities. A finite disc thickness weakens the surface gravity and lowers the critical value where the disc undergoes instability. For example, Goldreich & Lynden-Bell (1965a) showed that $Q_\mathrm{c} = 0.676$ for a single-component thick disc. In addition, the onset for *non-axissymmetry* occurs at higher values of Q_g, both for 2D ($Q_\mathrm{g} \sim 1.7$) and 3D discs. Extended stability analysis taking thickness and multiple components (collisional and/or collisionless has developed by e.g. Jog & Solomon (1984); Romeo (1992); Rafikov (2001).

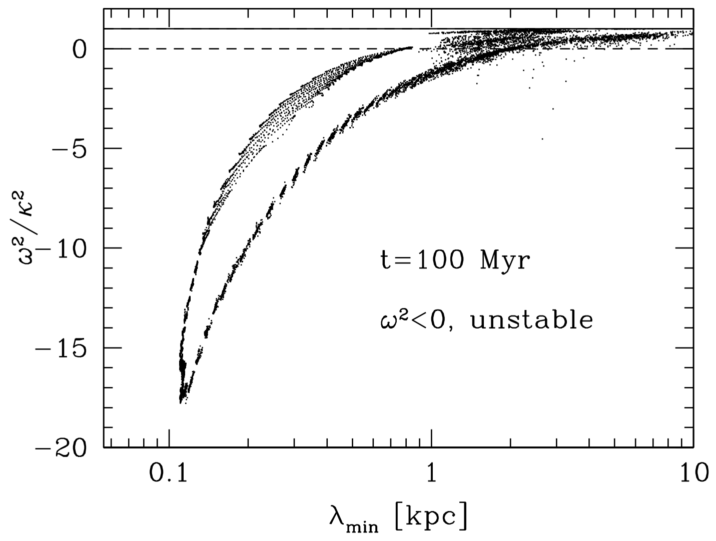

Figure 4.9: Most unstable wave lengths for the whole disc around time of fragmentation i.e. $t \sim 100\,\mathrm{Myr}$. A large part of the disc is unstable at $\sim 100\,\mathrm{pc} - 2\,\mathrm{kpc}$, scales that will collapse to the initial cloud distribution.

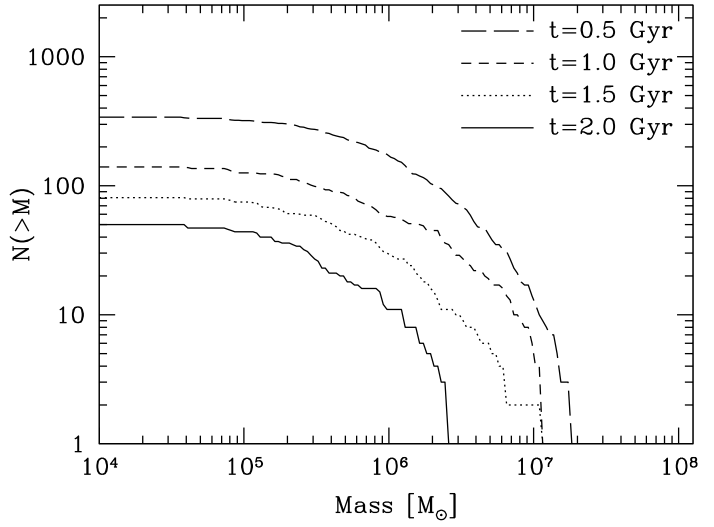

Figure 4.10: Cumulative mass spectrum of individual "molecular" clouds ($n > 100\,\mathrm{cm}^{-3}$) in RUN3.

In this section we mainly focus on the more unstable gas component as it will be the main driver of turbulence compared to the stellar component which shows a higher degree of stability at all times ($Q_* > 2$).

We start by investigating the early time evolution when the initial cloud population forms. To do this we need to quantify the most unstable length scales. The dispersion relation for axisymmetric disturbances

$$\omega^2 = \kappa^2 - 2\pi G \Sigma_{\rm g} k + \sigma_{\rm eff}^2 k^2, \tag{4.9}$$

where ω is the growth rate and k is the wavenumber of the perturbation. Instability demands that $w^2 < 0$ and the most unstable mode is simply the minima of Eq. 4.9, i.e.

$$\lambda_{\rm min} = \frac{2\sigma_{\rm eff}^2}{G\Sigma_{\rm g}}. \tag{4.10}$$

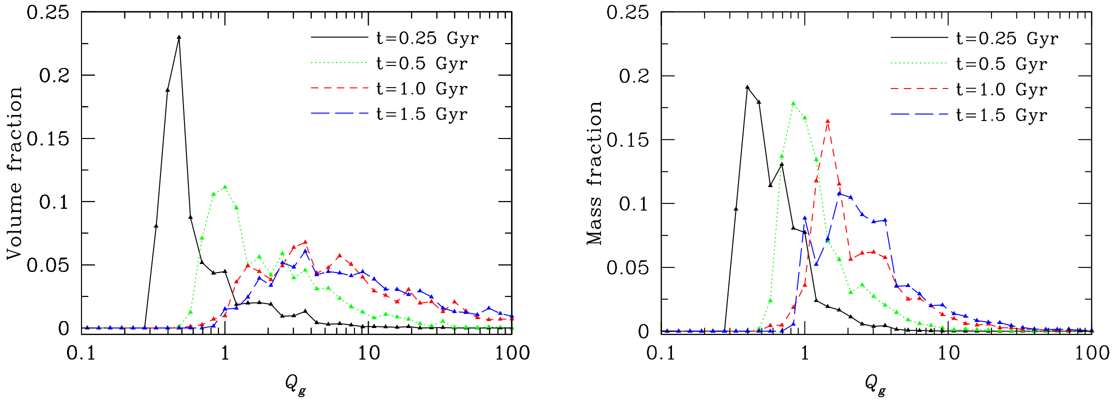

Figure 4.11: (*Left*) Volume distribution of $Q_{\rm g}$ values for the whole disc in RUN1. The low $Q_{\rm g}$ values at early times indicates the formation and existence of clouds. Later times shows an equilibrium distribution that changes little with time. (*Right*) Mass distribution of $Q_{\rm g}$ values. Most of the disc is distributed over $1 < Q_{\rm g} < 4$.

We calculate all components the dispersion relation, and $Q_{\rm g}$, on a polar grid with subregions $30\,{\rm kpc}/\Delta R = 40$ and $2\pi/\Delta\theta = 60$. Fig. 4.9 shows the minimum of the dispersion relation and their related growth factors at $t = 100\,$Myr. The bimodal distribution of points might seem odd but is merely a reflection of the initial $Q(r)$ (see Fig. 4.2), where the same value of $Q_{\rm g}$ exists at different radii (and different densities) and hence show the same w^2 at different $\lambda_{\rm min}$. We note that scales of $100\,{\rm pc} < \lambda_{\rm min} < 2\,{\rm kpc}$ are in the unstable regime, where the smaller scales show larger negative values of ω^2. It is reassuring that smaller scales remain stable due to the imposed Jeans capturing EOS discussed in Sect. 4.2. These gravitational instabilities set the initial conditions for the clouds. The evolution of the cumulative mass spectrum, for "molecular" gas ($n > 100\,{\rm cm}^{-3}$), in RUN1 is shown in Fig. 4.10. We have calculated the mass spectrum by simply discerning individual pieces of high-density gas in the disc. This method is crude and occasionally overestimates the mass of clouds in the central parts of the discs where the gas density is high and crowding artificially identifies several clouds as one. Disregarding this, we note that the spectrum around $t \sim 1.5\,{\rm Gyr}$ occupies similar values as that of local group spirals (Blitz et al., 2007), where the most massive clouds are $\sim 10^7\,M_\odot$. The small mass truncation is due to limited resolution. As the simulations lack important small scale physics, e.g. MHD, radiative transfer, cosmic rays etc., the cloud population is long-lived and only reflects a true ISM *in a statistical sense*. This notion should not be a problem for the source of turbulent velocity dispersions that, as shown in Sect. 4.3, is due to large scale gravitational drags that most probably is independent of the small scale gas state close to or inside of the cloud complexes.

We now turn to the subsequent evolution. Fig. 4.11 shows the time evolution of the distribution of $Q_{\rm g}$ values for the whole disc. The left panel shows the volume fraction that different values of $Q_{\rm g}$ occupy while the right panel treats the mass fraction. This figure illustrates the complexity of the simulated discs and why azimuthally averaged $Q_{\rm g}(r)$ can be misleading. Initially, the disc shows a low spread of $Q_{\rm g}$ around a value of a few. As the disc cools down and undergoes gravitational instability (after $t \sim 0.1\,{\rm Gyr}$) this simple picture changes. At $t = 0.25$ Gyr, the disc has undergone fragmentation and the distribution is confined to $0.2 < Q_{\rm g} < 1$. The peak of the distribution, and the dispersion, gets larger with time. Part of this owes to star formation that acts to lower $\Sigma_{\rm g}$. For $t > 1.0$ Gyr the disc evolves into what appears to be an equilibrium state, spanning a large range in $Q_{\rm g}$-values ($0.5 \lesssim Q_{\rm g} \lesssim 10^2$). This co-existence of $Q_{\rm g}$-value in a patchy galactic disc is in agreement with the analysis of Wada et al. (2002) for their two-dimensional models. The mass fraction distribution follows a similar evolution, approximately reaching an equilibrium state after $t > 1.0$ Gyr. However, a significant part of the $Q_{\rm g}$ distribution at late times now populates the unstable or marginally stable values. At 1.5 Gyr, most of the disc is distributed around $1 \lesssim Q_{\rm g} \lesssim 4$. Regardless of the exact distribution, the dominating existence (by mass) of unstable and marginally stable regions of the disc is of great importance for generating a global gravitoturbulent state which we will return to in Sect. 4.3. Without the onset of gravitational instabilities, the gas would approximately stay on circular orbits.

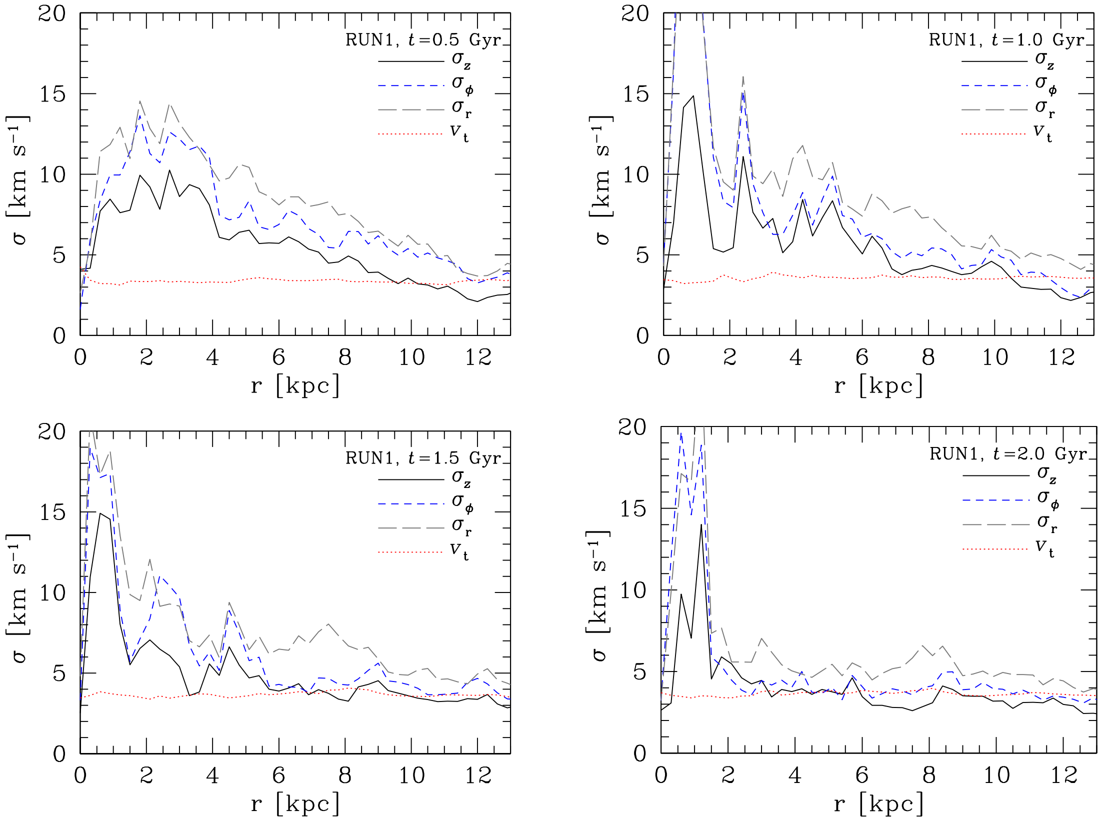

Figure 4.12: Velocity dispersions of the vertical (σ_z), angular (σ_ϕ), radial (σ_r) and thermal component (v_t) of the HI gas in RUN1 at $t = 0.5, 1.0, 1.5$ and $2.0\,$Gyr.

At late times the mass in the disc is dominated by the stellar component. To get an understanding of its influence on stability one can use an approximate stability parameter (Bertin & Romeo, 1988; Romeo, 1994) which in standard regimes is of the form

$$Q \approx Q_* \left(1 - 2\frac{\Sigma_\mathrm{g}}{\Sigma_*}\right). \tag{4.11}$$

At late times, the stellar component is in the range $1.5 \lesssim Q_* \lesssim 5$ in the star forming region which together with the values shown in Fig. 4.11 assures us that the multicomponent disc will never be completely stable, at least locally.

The Turbulent ISM — Having characterized the multiphase ISM in terms of phases and stability, we can now properly address the main topic of this work, the HI velocity dispersions.

☐ *Velocity Dispersions* — The observational tradition is to model HI profiles using one or multiple Gaussians where the flux is a function of the velocity v as

$$f(v) = \frac{1}{\sigma\sqrt{2}} \exp\left[-\left(\frac{1}{2\sigma^2}\right)(v - v_0)^2\right] \tag{4.12}$$

where v_0 is associated with the peak flux and σ is the actual velocity dispersion. Broadening of spectral lines is mainly due to thermal and Doppler/turbulent effects. We will discuss the thermal effects in terms of the thermal velocity $v_\mathrm{t} = \sqrt{RT/\mu}$ (i.e. the isothermal sound speed of the gas), where R is the gas constant and μ is the molecular weight. Random bulk motion of the gas is quantified in terms of its turbulent velocity dispersion σ_t. We calculate the net observable dispersion by adding the turbulent and thermal contribution in quadrature, i.e. $\sigma^2_\mathrm{eff} = v^2_\mathrm{t} + \sigma^2_t$.

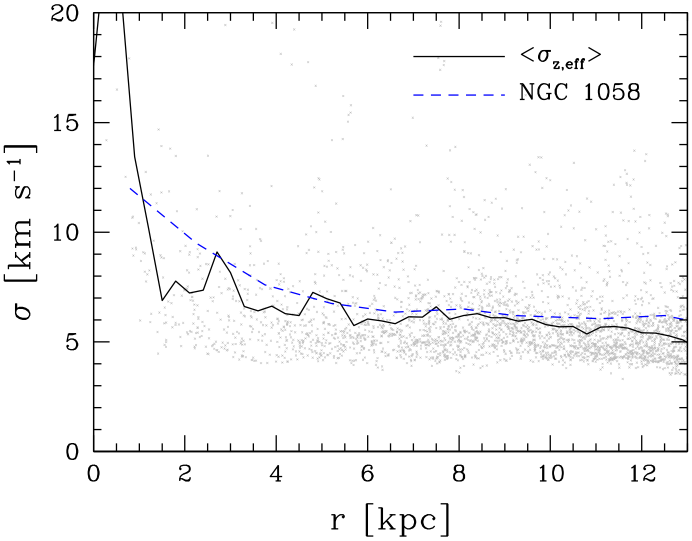

Figure 4.13: Observed velocity dispersion (blue dashed line) of the face on galaxy NGC 1058 compared to the effective vertical dispersion (black solid line) of RUN1 at $t = 1.5\,\mathrm{Gyr}$. Grey dots indicate locally measured dispersions, see text.

In the following analysis, we characterize only the gas that would be observed as HI and not the dense clouds that will consist of mainly molecular gas. We therefore use the criteria $\rho < 10\,\mathrm{cm}^{-3}$ (star formation threshold) and $800\,\mathrm{K} < T < 10\,000\,\mathrm{K}$. This choice is suitable as it is more likely to exist outside of the denser spiral arms as well as in the outer regions of the disc, which is where observations lack an explanation. The velocity dispersion is calculated by randomly sampling the galactic discs using synthetic observational patches ($5\,000$ patches were used for the data described here) of size $\sim 700\,\mathrm{pc}$. We choose this size as this is the stated scale below which bulk motions are expected to be responsible for the observed dispersions (Petric & Rupen, 2007). In each patch we calculate both the mass weighted turbulent velocity dispersions σ_t and the mass weighted mean thermal velocity. Weighting by mass is well motivated as HI emission is strongly correlated with the local density.

The panels in Fig. 4.12 shows a time evolution of the radial behaviour of the velocity dispersion for RUN1, where σ_z is the vertical dispersion component, σ_r the radial and σ_ϕ the angular. σ_z show typical values of $\sim 15\,\mathrm{km\,s}^{-1}$ in the center and declines to $\sim 3-5\,\mathrm{km\,s}^{-1}$ at large radii. The velocity dispersion is clearly anisotropic as $\sigma_r > \sigma_\phi > \sigma_z$ at all times. It is interesting that the ratio of the dispersions roughly follow the epicyclic predictions for a collision-less system, i.e. $\sigma_r = 2\Omega\sigma_\phi/\kappa$. A similar result was found by Bottema (2003). The planar dispersion σ_{xy}, i.e. the RMS value of the radial and angular dispersions, is a factor of ~ 2 larger than the vertical dispersion at all times and radii. The thermal component of the gas lies in the range $3-5\,\mathrm{km\,s}^{-1}$ in agreement with a warm gas component ($T \sim 1000-2000\,\mathrm{K}$). The planar velocity dispersion is supersonic or transsonic at all times and radii while the vertical dispersion is generally transsonic, turning sub-sonic at large radii. This means that the thermal component becomes as important as the turbulent at large radii for the total observable velocity dispersion. By considering a minimal observable (σ_eff) for the z-component, we clearly find an agreement with the observed HI dispersions values described in Sect. 4.1 (i.e. $\sigma_\mathrm{eff} \sim 12-15\ \mathrm{km\,s}^{-1}$ in the inner parts declining to $\sim 4-6\ \mathrm{km\,s}^{-1}$ in the outer). Any inclination would boost these values due to the σ-anisotropy. The same analysis has been performed on the higher resolution simulation RUN2 with no significant difference in the results.

We now directly compare our simulations to the HI data of the spiral galaxy NGC 1058 (Dickey et al., 1990; Petric & Rupen, 2007). NGC 1058 is a suitable object for comparisons, as it is comparable in size, surface density and peak rotational velocity (derived to be $\sim 150\,\mathrm{km\,s}^{-1}$) to our simulated disc. The galaxy also has a low star formation rate, SFR $\sim 3.5 \times 10^{-2}\,M_\odot\,\mathrm{yr}^{-1}$ (Ferguson et al., 1998), which places it in the flat part of Fig. 4.1. Furthermore, by being an almost perfectly face-on galaxy (inclination of $4-11\,°$), we can disentangle the vertical component from the planar. In Fig. 4.13 we compare $\sigma_{z,\mathrm{eff}}$ of RUN1 at $t = 1.5\,\mathrm{Gyr}$

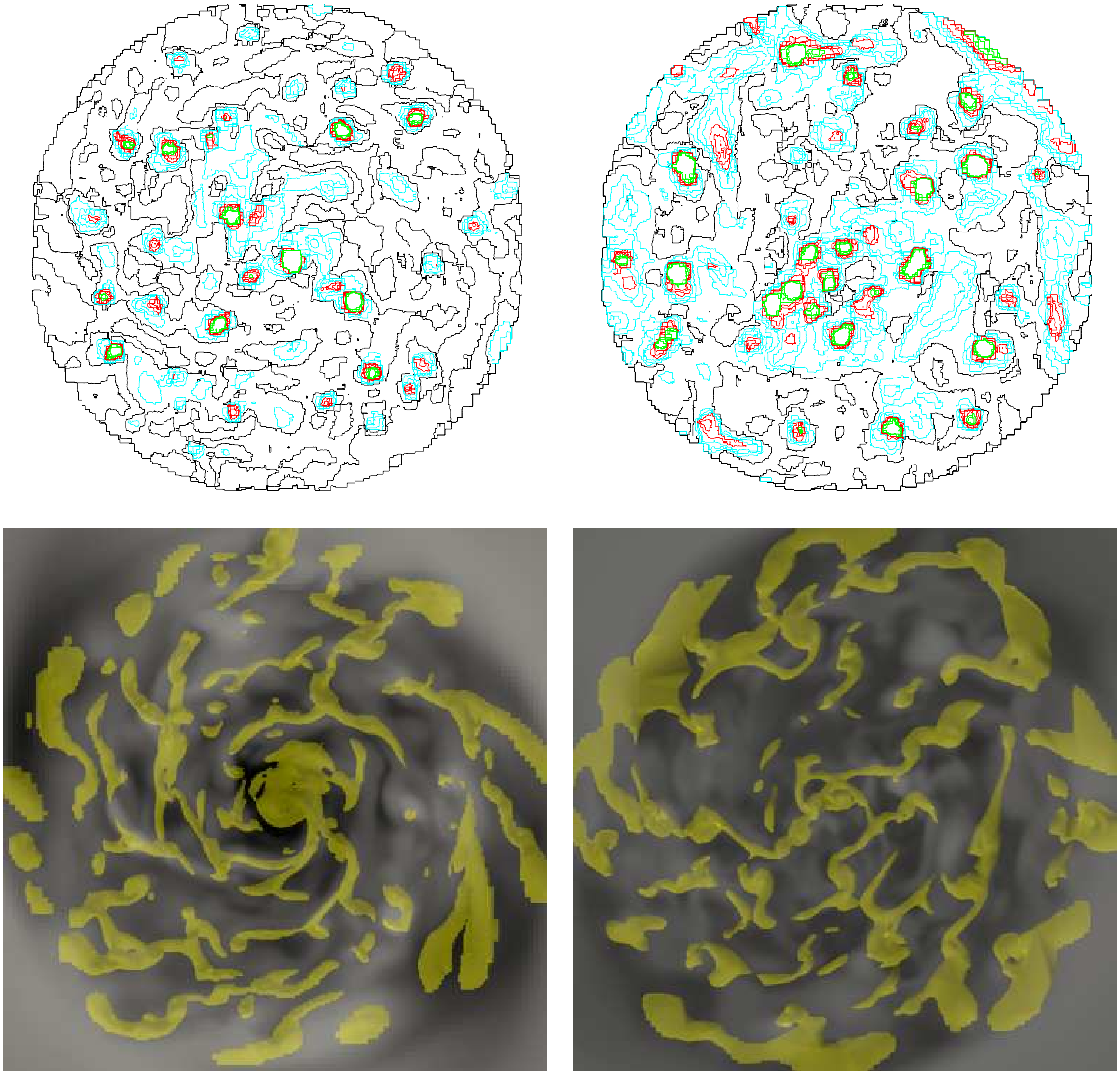

Figure 4.14: Distribution of vertical velocity dispersions in RUN1 (*top left*) and RUN3 (*top right*) at $t = 2.0\,\mathrm{Gyr}$ calculated for HI only (see text). The plotted region is $30\,\mathrm{kpc}$ across. The contour levels are in $\mathrm{km\,s^{-1}}$ in steps of $1\,\mathrm{km\,s^{-1}}$. Black is used between 3 and 5, cyan for 6 to 8, red for 9 to 11 and green for 12 to 14 $\mathrm{km\,s^{-1}}$. The bottom panels show contours of the gas with $n > 0.2\,\mathrm{cm^{-3}}$ of the same regions but slightly zoomed out ($40\,\mathrm{kpc}$ across).

with the observational data of NGC 1058. Our simulation not only reproduces the magnitude of the velocity dispersion but also the declining radial shape.

The spatial distribution of the vertical velocity dispersions in NGC 1058 is very patchy with several peaks of $\sigma > 10\,\mathrm{km\,s^{-1}}$ (see Fig. 5 of Petric & Rupen (2007)). This observation is reproduced by our simulations, as shown in Fig. 4.14 where we plot the contours of $\sigma_{\mathrm{eff},z}$ in RUN1 and RUN3 at $t = 2.0\,\mathrm{Gyr}$ as well as the corresponding density fields. The high density gas is distributed in a flocculant spiral structure, reminiscent of the HI observation of M33 (Deul & van der Hulst, 1987; Engargiola et al., 2003). Analyzing the simulation at a late epoch is preferred as the mass spectrum of dense clouds has evolved to a rather realistic state, see Fig. 4.10. The strongest peaks ($\sigma_{\mathrm{eff},z} > 10\,\mathrm{km\,s^{-1}}$) are associated with dense clouds while mildly turbulent regions (cyan levels at $\sigma_{\mathrm{eff},z} \sim 6 - 8\,\mathrm{km\,s^{-1}}$) often exist in inter-cloud/arm regions of strong shear. Regions of large velocity dispersion related to clouds also extend several kpc away from their radius of influence. We note that RUN3 even at late times displays a few $\mathrm{km\,s^{-1}}$ larger velocity dispersions in diffuse regions,

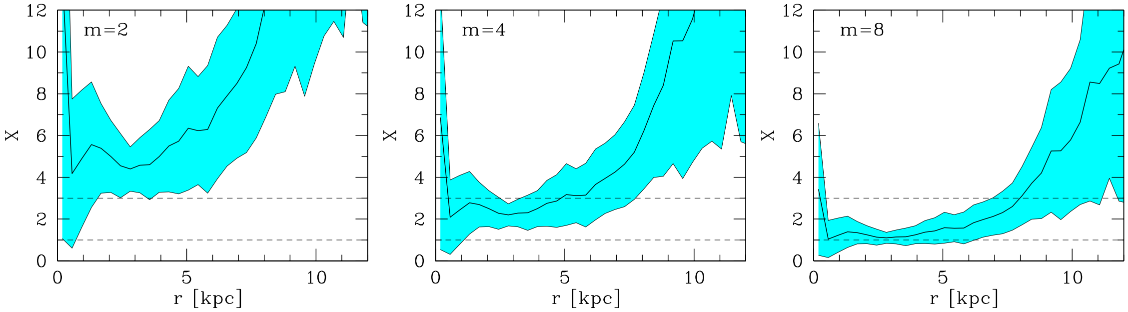

Figure 4.15: Radial dependance of the X-parameter for $m = 2, 4$ and 8. Higher order modes ($m \gtrsim 4$) are in the range $1 < X < 3$ where swing amplification is efficient.

probably due the more prominent warm gas phase.

☐ *What is the Driver of ISM Turbulence?* The drivers of the turbulent component of the velocity dispersion are gravity and shear. In Sect. 4.3 we found that the galactic discs have, by mass, a wide spectrum of $Q_{\rm g}$ values where a significant part sits at local marginal stability for a finite thickness disc. The 2D shearing box simulations by Kim & Ostriker (2007) showed that a marginally stable gas discs at $Q_{\rm g} \sim 1.2$ can generate velocity dispersions of the order of the local sound speed, decreasing for larger $Q_{\rm g}$-values (see their Fig. 12). The origin of turbulence was here attributed to swing amplification. Note that the 3D structure of our discs necessitates values $\sim 25\,\%$ lower for an equivalent stability. A full turbulent outcome of more unstable discs ($Q_{\rm g} < 1.2$) was not studied as the velocity field would then only be a response to very strong density inhomogeneities. Local shearing boxes are useful for understanding the mechanism that drives turbulence at *specific values* of $Q_{\rm g}$. As our simulations show a wide range of $Q_{\rm g}$ values we have the combined spectrum of swing amplified turbulence across the whole disc for gas that locally behaves in accordance with the simulations of Kim & Ostriker (2007) for $Q_{\rm g} > 1.2$. In a statistical sense there will always be regions with a $Q_{\rm g}$-value low enough to tap large velocity dispersions from the swing mechanism which is confirmed in Fig. 4.14 where intermediate values of $\sigma_{{\rm eff},z}$ is associated with waves. To quantify this it is useful to use the X-parameter (Toomre, 1981) defined as

$$X = \frac{k_{\rm crit} R}{m} \tag{4.13}$$

where $k_{\rm crit} = \kappa^2/2\pi G\Sigma$, $\Sigma = \Sigma_{\rm g} + \Sigma_*$ and m is the number of arms. It has been shown by Jog (1992) that swing-amplification is very effective in the gas component in multicomponent discs. Even when both the gas and stars separately are stable ($Q_{\rm g} = Q_* = 2$), their gravitational coupling can amplify waves in he gas for values of X not much larger than unity. In one component, marginal stability and $1 < X < 3$ can be considered sufficient to assure amplification (Toomre, 1981). Fig. 4.15 shows the radial dependance of the X-parameter for $m = 2, 4$ and 8 at $t = 2.0$ Gyr. We see that amplification is efficient for $m \geq 4$ which confirms the high-order flocculant spiral structure in Fig. 4.14.

For small values of $Q_{\rm g}$, where the gas locally has undergone full non-linear gravitational instability, the situation is different. The cold phase dominates the gas mass, even at early times and is therefore locally the most important gravitational source. Direct cloud merging and tidal interactions stirs the inter-cloud medium both radially and vertically. Apart from stirring the gas, the clouds also dissipate energy thermally in shocks which regulates the warm phase of the ISM, forming the $\sim 4 - 5\,{\rm km\,s^{-1}}$ thermal components of $\sigma_{\rm eff}$. We observe cloud formation, merging, scattering and reformation during the whole simulation time. Formation in a shearing environment causes dense structures that are not tightly bound to the actual clouds to stretch into waves and filaments. This triggering of wave-like perturbations, and its associated irregular velocity field in the ISM, is a key role of the clouds which was realized already by Julian & Toomre (1966). As for the marginally stable gas surrounding the gas, the leading waves swing and amplify, inducing gravitational torques in the gas and hence increasing the local velocity dispersion (e.g. Kim et al., 2002; Kim & Ostriker, 2007). Fig. 4.16 shows a typical patch of the disc in RUN1 at $t = 1$ Gyr, confirming this notion. We note that

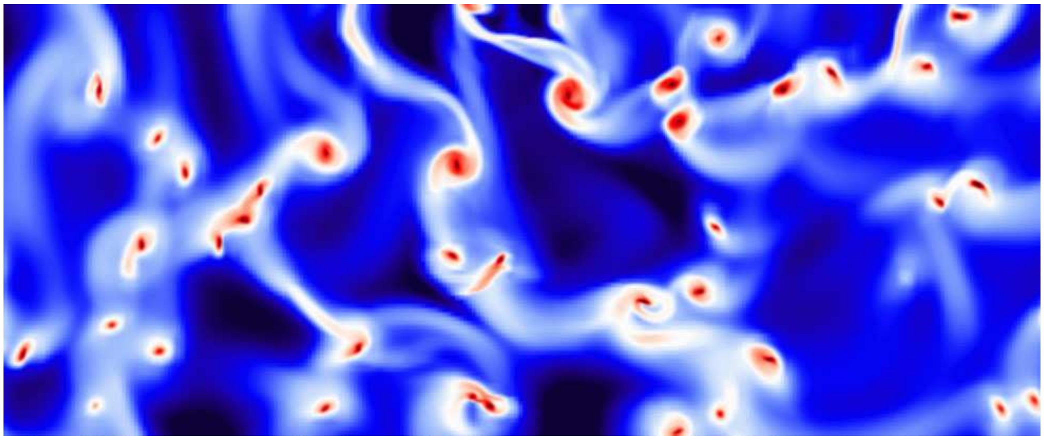

Figure 4.16: Density plot of a $24 \times 10\,\mathrm{kpc}^2$ region centered over $\{x, y, z\} = \{0, -5.0, 0\}\,\mathrm{kpc}$ in RUN1 at $t = 1.0\,\mathrm{Gyr}$. The colour map is here chosen to enhance the visual appearance of the clouds and filaments. The galactic rotation is here clock-wise. Filamentary structures is always associated with the clouds. Also, all clouds excite waves, many of them leading which will swing into trailing ones.

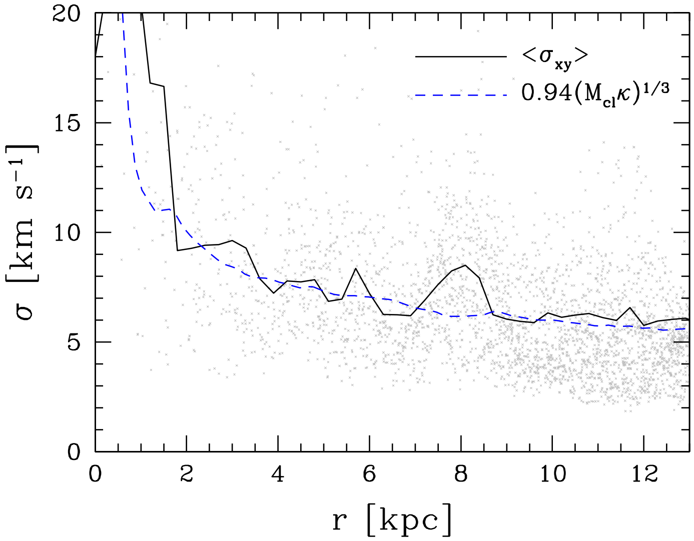

Figure 4.17: Planar velocity dispersion of RUN1 at $t = 2.0\,\mathrm{Gyr}$ (black solid line) compared to the relation derived by Gammie et al. (1991) (blue dashed line) for cloud scattering. The epicyclic frequency is obtained from the simulation and the cloud mass is chosen to be $3.5 \times 10^6 M_\odot$.

this processes is analogous to the energy extraction from background shear at a rate $T_{R\phi} d\Omega/d\ln R$, where the $T_{R\phi}$ stress tensor includes the contribution from Reynolds and Newtonian stresses, to induce local velocity dispersion as outlined by Sellwood & Balbus (1999).

But how can we quantify the impact of the cloud motions? Let us assume that the motions of the cloud ensemble are representative of the turbulent ISM. The swing amplifier might play a very fundamental role for turbulence but as the clouds effectively trace the large scale waves and constitute the majority of the mass, the assumption that turbulence is associated with cloud motions is a fairly good approximation. Cloud-cloud interaction can be modelled as gravitational scattering and has been studied analytically by e.g. Jog & Ostriker (1988) and Gammie et al. (1991). The semi-analytical perturbation theory model of Gammie et al.

(1991) predicts a planar velocity dispersion

$$\sigma_{xy} \approx 0.94(GM_{\rm cl}\kappa)^{1/3}, \tag{4.14}$$

where $M_{\rm cl}$ is a typical mass of a cloud. This relation is derived for a two-dimensional, two body encounter on radially separated orbits in a shearing disc. However, as these clouds are the main local perturbers by mass we can assume that any diffuse HI gas will approximately be dictated by the cloud ensemble velocities. In Fig. 4.17 we plot σ_{xy} for RUN1 at $t = 2.0\,{\rm Gyr}$ against Eq. 4.14 using a cloud mass $M_{\rm cl} \approx 3.5 \times 10^6 M_\odot$, and $\kappa(r)$ of the gas in the simulation. We stress that $M_{\rm cl}$ is in the high-end of a typical GMC mass spectrum (Blitz et al., 2007). As the clouds in our simulations are submerged in massive HI envelopes, the largest clouds are closer in mass to that of GMC complexes, GMAs (Giant Molecular Associations) or super-clouds (Rosolowsky et al., 2007). A more realistic analysis should include the full spectrum of cloud masses and their radial distribution but even this simple analysis renders a good agreement with the measured dispersions. The weak dependence on κ can explain why most non star-bursting galaxies seem to plateau at a velocity dispersion between 7 and $11\,{\rm km\,s^{-1}}$ (see Fig. 4.1 and discussion in Sect. 4.1). The rotational velocity varies from $\sim 100\,{\rm km\,s^{-1}}$ to $\sim 300\,{\rm km\,s^{-1}}$ for most spirals. By assuming a flat rotation curve, $\kappa \propto v_{\rm c}$, the actual change in cloud velocity dispersion between the two limits is, assuming an invariant cloud spectrum, only by a factor of $3^{1/3}$ ($\approx 44\,\%$).

To conclude, the full picture of gravity driven turbulence in the ISM is based on the existence of dense clouds and filamentary structures all adding to the turbulence budget. Marginally stable gas and cloud-induced filaments generate turbulence through gravitational torques from the swing-amplifier. Clouds also scatter gravitationally, perturbing the local velocity field even further as well as shock-heating the ISM. The source of turbulence in both cases is self-gravity coupled with shear that in turn converts ordered circular motion of the gas to random velocities, hence tapping rotational energy from the disc. In a sense, self-gravity can here be regarded as a form of viscosity (Jog & Ostriker, 1988). We argue that the mechanisms described in this section serve as a baseline level of turbulence for galaxies where any excess observed velocity dispersion is caused by additional sources such as supernovae activity or magnetic fields coupled with shear. Finally, we note that many of the disc characteristics, e.g. σ and $Q_{\rm g}$, show signs of reaching a statistical equilibrium state at late times in the simulations. This is an indication that there is a constant supply rate of energy to the system, here coming from galactic rotation, maintaining a constant level of turbulence. However, on close inspection these quantities are slowly decaying (e.g. the mean velocity dispersion is slowly decreasing, see Fig. 4.12) due to star formation depleting the disc of gas. Inclusion of realistic gas accretion would affect the temporal evolution of the latter quantities.

□ *Varying the Baryon Fraction, Shearing Motions and Cooling Floor* — In order to explore the effect of the gas density in the initial disc, we keep all parameters fixed apart from the disc mass and carry out an additional simulation (RUN6) in which the gas mass is one third of our fiducial value. This changes the surface density profile and thus the strength of self-gravity. This disc is extremely light and is unstable only within $\sim 8\,{\rm kpc}$, leading to an abundance and mass spectrum of cold gas clouds that is significantly smaller than in RUN0 or RUN1 as seen in Fig. 4.18. Clouds are here defined as isolated clumps of gas satisfying $n > 100\,{\rm cm^{-3}}$. The vertical shift, when comparing RUN0 and RUN6, is mostly due to the larger area of instability in the latter simulation and the RUN0 and RUN1 off-set is simply due to star-formation acting to deplete the clouds of high-density gas. We can quantify the observed mass-shift using the arguments presented by e.g. Escala & Larson (2008) where the high-mass end of the spectrum is set by the largest modes not to be stabilized by shear, i.e. the maximum cloud mass will be set by

$$M_{\rm cl}^{\rm max} = \frac{\pi^4 G^2 \Sigma_{\rm gas}^3}{4\Omega^4}. \tag{4.15}$$

A factor of 3 decrease in $\Sigma_{\rm gas}$ therefor leads to $M_{\rm cl}^{\rm max}$ being 27 times smaller which is in excellent agreement with the mass spectrum shift in Fig. 4.18. The most massive clouds in RUN6 are $\sim 10^6 M_\odot$, leading to weaker cloud-cloud encounters and swing amplified waves, directly lowering the accelerations imparted on nearby gas parcels. This leads to a lower overall velocity dispersions by a factor proportional to $M_{\rm cl}^{1/3} \sim \Sigma$ (from Eq. 4.14 and Eq. 4.15). An off-set by a factor of ~ 3 is confirmed in Fig. 4.19 which shows the ratios of $\sigma_{xy}(r)$ of the different simulations.

To isolate the effect of the shearing motions we keep all parameters the same as our fiducial RUN1, apart from the rotation curve set by the dark matter halo which is constructed to be much flatter (RUN8). This is

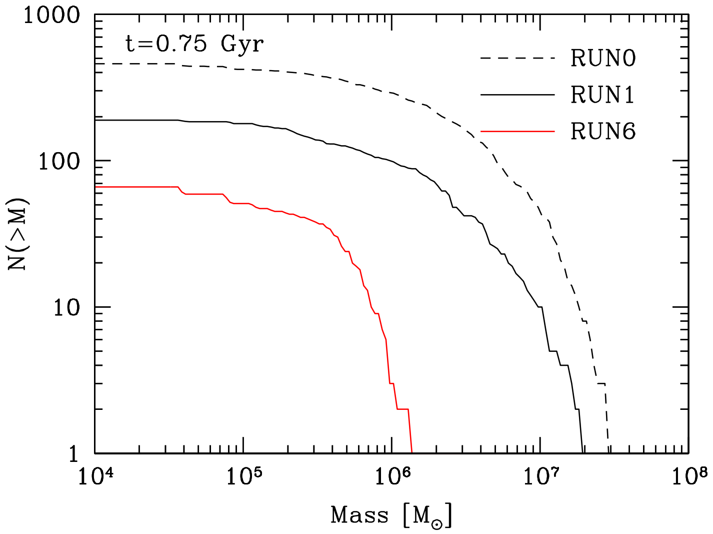

Figure 4.18: Cumulative mass spectrum of clouds in RUN0 (dashed line), RUN1 (black) and RUN6 (red) at $t = 0.75$ Gyr. The smaller region of instability in the less dense disc in RUN6 brings down the total number of clouds and the mass offset is related to a decrease of the largest wavelength to be stabilized by shear, see text for discussion.

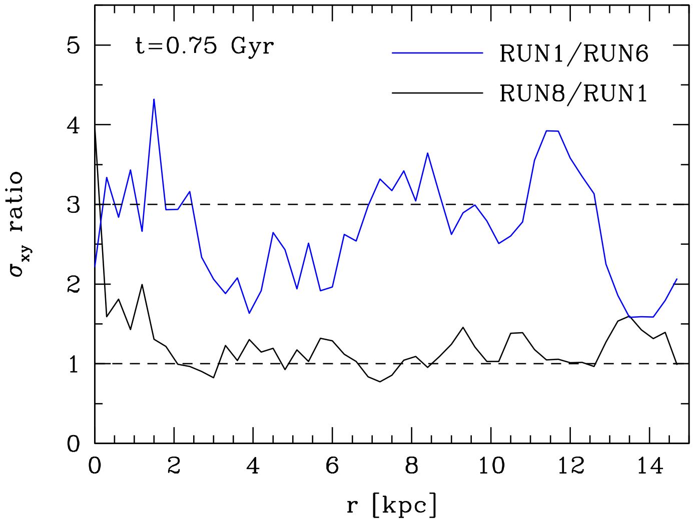

Figure 4.19: Ratios of planar velocity dispersions for RUN1, RUN6 and RUN8. The RUN1/RUN6 ratio illustrates the impact of the surface density of the gas, where the more massive disc in RUN1 has ~ 3 times larger velocity dispersion. The increased shear in RUN8 increases the velocity dispersion when compared to RUN1, but only significantly in the central parts of the disc.

achieved by increasing the concentration parameter to $c = 40$ whilst maintaining the peak circular velocity to be similar to the rest of our simulations. The swing instability and turbulence induced by shearing motions should be stronger since the epicyclic frequency, defined as

$$\kappa^2 = [r\frac{d\Omega^2}{dr} + 4\Omega^2], \tag{4.16}$$

where $\Omega = v_c/r$, increases. For a $v_c \sim \sqrt{r}$ this mean that $\kappa \sim 1/\sqrt{r}$ and for $v_c \sim$constant, $\kappa \sim 1/r$. The

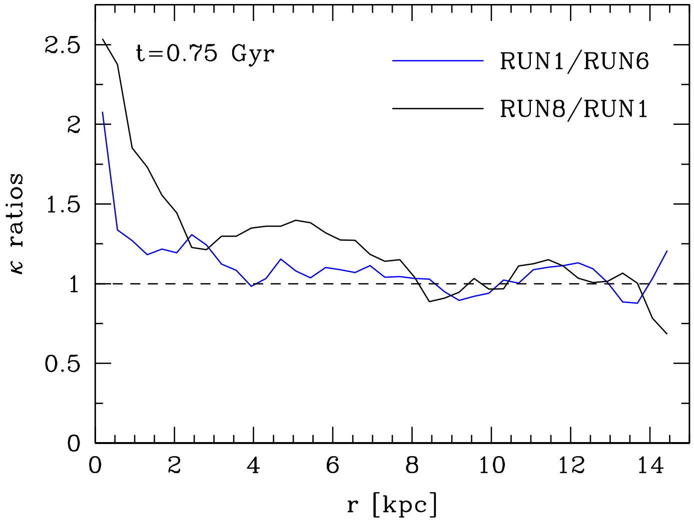

Figure 4.20: Ratios of epicyclic frequencies for RUN1, RUN6 and RUN8. The RUN1/RUN6 ratio is above unity in the central parts owing to excess mass transport to the center in RUN1 and hence an increase in central shear. The overall increase in RUN8, where the central parts of the disc are more affected, is expected from the change in dark matter halo concentration.

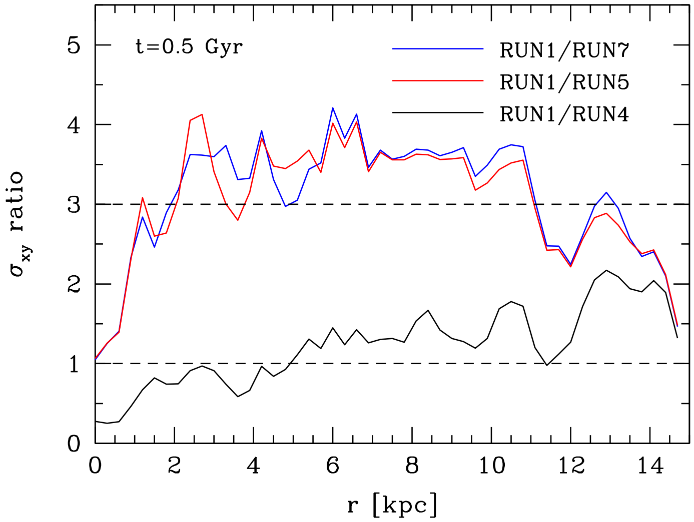

Figure 4.21: Ratios of planar velocity dispersions for RUN1, RUN4, RUN5 and RUN7 showing the effect of a temperature floor for the gas cooling. The velocity dispersion in RUN5 and RUN7 are lower by a factor of 3-4 compared to RUN1 due to the inability to undergo gravitational instability. The RUN4 disc is allowed to cool enough to capture the largest clouds forming, hence having similar velocity dispersions as RUN1.

former v_{c} roughly describes RUN1 and the latter RUN8. Fig. 4.20 shows $\kappa(r)$ for RUN1, RUN6 and RUN8 at $t = 0.75\,\mathrm{Gyr}$. RUN1 and RUN6 show similar values apart from the inner part of the disc where turbulent viscosity in RUN1 has dragged in more mass compared to RUN6, rendering a more active region. RUN8 shows a steeper behavior, having a κ up to 2.5 times larger in the very center of the disc down to a ratio of unity at $r \sim 8\,\mathrm{kpc}$. The larger shear causes a larger $\sigma_{xy}(r)$ compared to RUN1, as seen in Fig. 4.19. The

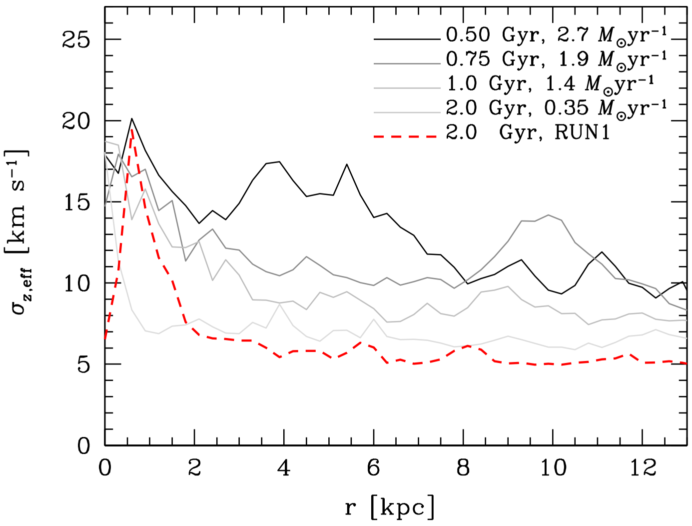

Figure 4.22: Effect of star formation rate on the observed vertical velocity dispersion in RUN3. The different lines indicate the values at different times and therefore also for different SFRs. There is a clear trend that a lower stellar activity lowers the measured dispersion, approaching the baseline observed dispersion given by RUN1 at $t = 2.0\,\mathrm{Gyr}$ (thick red dashed line).

effect is strong in the central parts of the disc ($1.5 - 2$ times larger) and for $r > 2\,\mathrm{kpc}$ the ratio is closer to unity.

It seems evident that gravitational instability is the important driver of velocity dispersions. To explore this further, we artificially truncate the ability to cool gas at different temperatures, hence effectively setting a floor for Q_g . In RUN4, RUN5 and RUN7 we introduce a cooling floor for the gas at 10^3, 10^4 and $5\,000$ K respectively. In Fig. 4.21 we compare the ratios of $\sigma_{xy}(r)$ at $t = 0.5\,\mathrm{Gyr}$ and it is evident that RUN5 and RUN7 only show small turbulent motions while RUN4 essentially is as turbulent as the fiducial RUN1. The physical parameters are the same in all simulations except for the temperature of the gas. As $Q_\mathrm{g} \sim \sigma \sim \sqrt{T}$ it is straightforward to see that RUN5 and RUN7, based on the initial $Q_\mathrm{g} \sim 2 - 3$ in Fig. 4.2, never acquires $Q_\mathrm{g} < 0.676$. Both simulations only show a weak development of spiral structure. We note that it is plausible that a stellar component would increase the turbulent motions of the discs, as found by Kim & Ostriker (2007). Also, the effective observable $\sigma_{z,\mathrm{eff}}$ in RUN5 and RUN7 is dominated by the thermal component as is at large radii close the values found in RUN1. The cooling in RUN4 can lower Q_g by a factor ~ 3.2, hence bringing $Q_\mathrm{g} < 0.676$ while being assisted by non-axissymmetric instabilities. While the smallest scales to be unstable differ in RUN1 and RUN4 due to different pressure support, the larger unstable wavelengths are the same, only limited by the same amount of shear. The presence of the larger, and more dynamically important, clouds and gravitational instabilities makes the velocity dispersion ratio closer to unity.

We conclude that the mass, and hence the surface density, of the disc has a significant impact on the generated turbulent velocity field and we roughly find that, provided the disc is gravitationally unstable, $\sigma \sim \Sigma$. We associate this to the weaker gravitational instabilities present in the disc, which can be seen from the cloud spectrum. The shear of the disc also affects the magnitude of the velocity dispersions but in a weaker fashion. This less strong effect of shear might originate from the $\sim \kappa^{1/3}$-relation (Eq. 4.14) for cloud velocity dispersion.

☐ *Effect of Supernovae Feedback* — Self-gravity driven turbulence may be important for galaxies with a low SFR/Area but cannot be the dominant driver behind the large velocity dispersions correlated with high star formation rates. Observations suggests that galaxies with a SFR/Area $\geq$ few $\times\, 10^{-3} M_\odot\,\mathrm{yr}^{-1}\,\mathrm{kpc}^{-2}$ show velocity dispersions of several $10\,\mathrm{km\,s}^{-1}$, see Fig. 4.1. Dib et al. (2006) showed that strong SN feedback could explain the transition into this range but were unable to explain the other end of the spectrum. It is plausible that the reason for this stems from their local shearing box approximation that does not take the full

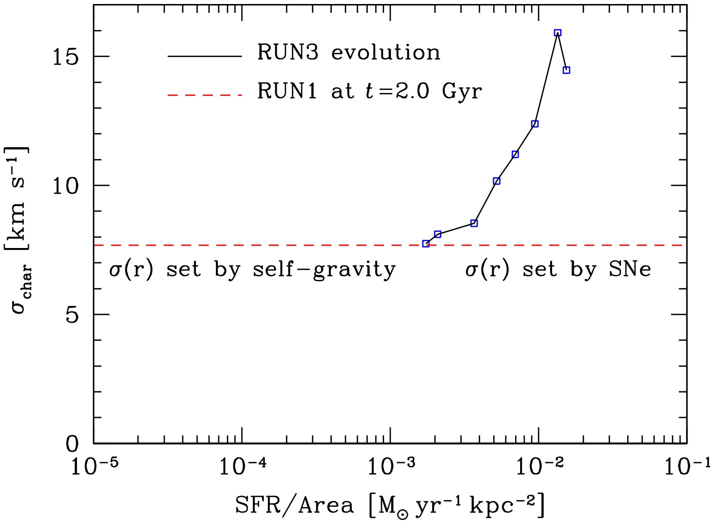

Figure 4.23: Effect of star formation rate on the observed vertical velocity dispersion in RUN3. The different lines indicate the values at different times and therefore also for different SFRs. There is a clear trend that a lower stellar activity lowers the measured dispersion, approaching the baseline observed dispersion given by RUN1 at $t = 2.0\,\mathrm{Gyr}$ (thick red dashed line).

disc dynamics into account. Furthermore, Dib et al. (2006) demonstrated that supernovae driven turbulence is sensitive to poorly known parameters such as efficiency, mass loading, timing etc. Due to computational cost, we can only study one set of parameters. However, as the star formation rate decreases with time (see Fig. 4.6) we are able to study its correlation with velocity dispersion. In Fig. 4.22 we plot the effective vertical dispersions for RUN3 at different times and hence different SFRs. The general amplitude of the dispersion declines with SFR and after $t = 1.5$ Gyr (SFR $\approx 0.74\,M_{\odot}\mathrm{yr}^{-1}$) there is little discrepancy between RUN1 and RUN3 suggesting that the effect of SN feedback has saturated. As seen in Fig 4.8, more warm gas (close to $\sim 10\,000\,\mathrm{K}$) exists in RUN3 explaining the $\sim 1\,\mathrm{km\,s}^{-1}$ off-set between RUN1 and RUN3 at large radii. This difference is also seen by inspection of the cyan contours in Fig. 4.14.

The data shown in Fig. 4.1 can be reproduced by averaging the velocity dispersion and SFR over a suitable area, hence obtaining a characteristic velocity dispersion $\sigma_{\mathrm{char}}(r)$. Dib et al. (2006) used the area $A = \pi(3r_0)^2$, where r_0 is the scale radius of the stellar disc. Fig. 4.23 shows the outcome of this procedure and we clearly detect a supernovae saturation to occur at a SFR/Area of $1 - 2 \times 10^{-3}\,M_{\odot}\,\mathrm{yr}^{-1}\,\mathrm{kpc}^{-2}$, where σ_{char} for RUN1 and RUN3 coincides at $t = 2.0\,\mathrm{Gyr}$. This transition from supernovae to self-gravity induced turbulence can explain why the velocity dispersion of NGC 1058 presented in Sect. 4.3 is in good agreement with our simulated disc. NGC 1058 has a derived SFR $\approx 3.5 \times 10^{-2}\,M_{\odot}\mathrm{yr}^{-1}$ (Petric & Rupen, 2007) which sets the SFR/Area well below $10^{-3} M_{\odot}\mathrm{yr}^{-1}\,\mathrm{kpc}^{-2}$ and hence into the regime where self-gravity induced turbulence can explain the observations.

4.4. Conclusions and Discussion

Three-dimensional, high-resolution hydrodynamical simulations using realistic modeling of star formation and evolution, show that a turbulent ISM naturally develops due to the coupling between gravitational instability and shearing motions. A multiphase medium develops in which cold dense clouds and filaments co-exists with a diffuse warm gas. When supernovae feedback is implemented, a hot phase is present. The marginally stable gas undergoes swing-amplification which both acts to amplify the local density as well as inducing gravitational torques. Cold and dense clouds undergo gravitational scattering, merging and tidal encounters. They also induce waves and filaments in the more diffuse gas which pumps energy into the turbulent process. The former mechanism stirs the gas even further and we note that the velocity dispersion of the clouds is a fairly good tracer of the HI velocity dispersion.

We summarize our main conclusions here:

- Gravitational instabilities in galactic discs leads to a population of massive cold clouds that undergo mutual gravitational interactions and merging. This cloud-cloud harassment process strips material and stirs the ISM. Both cloud interaction and the global non-axisymmetric instability of the disc create non-circular motions from initial ordered rotation. Waves and filaments are generated in the ISM which in turn swing-amplifies to generate further turbulent motions.
- Below a star-formation rate per unit area of $10^{-3} M_{\odot} \mathrm{yr}^{-1}$ we find that gravity alone can provide the energy source for maintaining the observed level of turbulence in the ISM of galaxies. The turbulent velocities in our M33 model galaxy have a mean value of $\sim 10\,\mathrm{km\,s^{-1}}$. By calculating an observable HI velocity dispersion, i.e. the contribution from both the turbulent and thermal components, we show that both the magnitude and radial profile is in good agreement of high-resolution HI surveys of e.g. NGC 1058 (Dickey et al., 1990; Petric & Rupen, 2007). In addition, we reproduce the observed patchy velocity dispersion map.
- Once the star-formation rate exceeds this value, supernovae feedback becomes the dominant driver of turbulence and the velocity dispersion increases with the star-formation rate. This agrees well with the general trend found by Dib et al. (2006).
- Lowering the initial gas density weakens the strength of gravitational instability and lowers the resulting cloud mass spectrum, which in turn leads to a lower disc velocity dispersion by a factor $\propto M_{\mathrm{cl}}^{1/3}$, as expected from a model in which self-gravity generates significant turbulent motions.
- A direct prediction of this scenario is that galaxies with lower gas fractions at a fixed halo mass should have lower velocity dispersions and different mass fractions in cold, warm and hot phases. Although, detecting the dependence on surface density is complicated by the fact that lower mass galaxies have a higher gas fraction in their discs (McGaugh, 2005). In addition, the reaction in low-mass systems to mild stellar activity has not been tested in this work and the outcoming HI velocity dispersion might conspire to render the plateau in Fig. 4.1.

It is important to note that these results do not rule out the importance of other contributing mechanisms such as supernova feedback or MHD processes, but underscore that self-gravity alone is an important, non-negligible source of turbulence in galactic discs. We believe that this work is complementary to alternative sources of turbulence, see Sect. 4.1. For example, Hennebelle & Audit (2007) considered turbulence driven by colliding flows in thermally unstable gas on very small (parsec) scales which are far from resolved in our simulations as we have aimed to resolve the large scale contribution from self-gravity that still would be within the large beam size ($\sim 700\,\mathrm{pc}$).

Other studies of large scale galactic turbulence includes Wada et al. (2002) and Wada & Norman (2007) who used an Eulerian code to simulate the dense central part of a galactic disc, where the cold molecular gas phase is dominating. Their results are in agreement with that found here, showing a complicated ISM with a wide range of Q-values. Using SPH, Gerritsen & Icke (1997) studied star formation and global evolution of the gas in a disc similar to NGC 6503. They demonstrated that a transient flocculant spiral structure with cold cloud complexes is naturally produced in the cold gas, in agreement with our results. The larger amount of warm gas was attributed to heating from stellar photons which is neglected in our work. The subsequent work by Bottema (2003) extended parameter space to understand the relationship between disc mass and global spiral structure and pointed out the success of swing amplification in predicting this. The measured gas velocity dispersion is similar to that obtained in this paper but was attributed to mechanical forcing from supernovae feedback.

Future work attempting more realistic formation of molecular clouds requires higher resolution and more sophisticated modeling of radiative physics and feedback in order to recover their full range of sizes, masses and life-times, which are affected by internal turbulence and strong feedback disruption. Even if the actual life and reformation times change, we believe that the global evolution of self-gravity driven turbulence will remain intact as it is not the absolute small scale state of the gas that governs the drag of the diffuse gas but the existence of massive interacting agglomerations. These massive clouds form through gravitational instability that requires seed fluctuations that may be triggered initially by numerical noise, but due to their rapid growth we expect the long term statistical behaviour to be representative. Similarly, our treatment of feedback is quite simplistic and could affect the lifetimes of the smaller mass clouds at the limit of our resolution, cf. end-state of RUN1 and RUN3 in Fig. 4.3, although we expect the larger clouds to be stable against these effects. Our initial conditions represent nearby well observed Sc galaxies such as M33 or NGC 1058. As we initialize the baryonic component as gas only we form very massive clouds at early times. However, these structure are significantly reduced in mass at late times due to star formation.

It is important to point out that the performed simulations do not include an old stellar population in the initial condition. This might change the global evolution to some extent and render a more pronounced spiral structure such as an $m = 2$ mode. Such a setup is rather complex and would involve a much larger parameter study. We have postponed this to a further study. In addition, we do not include a background UV field (far and local field). This will change the heating/cooling budget to some extent and can affect the gas density distribution. The global dynamics should however remain the same (see e.g. Wada et al., 2002).

Acknowledgments

O. Agertz would like to thank Andreas Burkert and Woong-Tae Kim for valuable discussion. We thank Mordecai-Mark Mac Low for valuable comments. We thank Doug Potter for making it possible to run the simulations on the zBox2 and zBox3 supercomputers (http://www.zbox2.org) at the University of Zürich.

5.
Disk Formation and the Origin of Clumpy Galaxies at High Redshift[26]

Observations of high redshift galaxies have revealed a multitude of large clumpy rapidly star-forming galaxies. Their formation scenario and their link to present day spirals is still unknown. In this *Letter* we perform adaptive mesh refinement simulations of disk formation in a cosmological context that are unrivalled in terms of mass and spatial resolution. We find that the so called ‘chain-galaxies’ and ‘clump-clusters’ are a natural outcome of early epochs of enhanced gas accretion from cold dense streams as well as tidally and ram-pressured stripped material from minor mergers and satellites. Through interaction with the hot halo gas, this freshly accreted cold gas settles into a large disk-like system, not necessarily aligned to an older stellar component, that undergoes fragmentation and subsequent star formation, forming large clumps in the mass range $10^7 - 10^9\,M_\odot$. Galaxy formation is a complex process at this important epoch when most of the central baryons are being acquired through a range of different mechanisms - we highlight that a rapid mass loading epoch is required to fuel the fragmentation taking place in the massive arms in the outskirts of extended disks, an accretion mode that occurs naturally in the hierarchical assembly process at early epochs.

5.1. Introduction

The morphology and star formation properties of high redshift galaxies are very different from present day quiescent spirals and ellipticals. Large clumpy irregular disks with kpc-sized star forming clumps as massive as $M_{\rm cl} \sim 10^7 - 10^9\,M_\odot$ are observed in the Hubble Ultra Deep Field (UDF) (e.g. Elmegreen et al., 2007, 2009), a population that is very rare today. ‘Chain galaxies’, first identified by Cowie et al. (1995), are believed to be high-redshift disky galaxies seen edge-on, while ‘clump cluster’ galaxies are their face on counterparts (Dalcanton & Shectman, 1996; Elmegreen et al., 2004). In optically selected samples, high redshift galaxies show very high star formation rates up to $100 - 200\,M_\odot{\rm yr}^{-1}$ (Daddi et al., 2004) and in recent spectroscopic observations they appear to be extended, though perturbed, rotating disks (Förster Schreiber et al., 2006; Genzel et al., 2006, 2008). The origin of these galaxies and how they connect and possibly evolve into present day spirals is still unknown. Gas rich major mergers give rise to large, bulge-dominated rotating disks (Robertson & Bullock, 2008) even though massive clumps can form at large radii, from globular clusters (Bournaud et al., 2008) to tidal dwarf galaxies (Elmegreen et al., 1993; Barnes & Hernquist, 1992). However, major mergers are not frequent enough (Dekel et al., 2009) and are more likely to be the origin of the rare, extremely high star forming, sub-millimeter galaxies (Zheng et al., 2004; Jogee et al., 2008).

Observational evidence (Elmegreen & Elmegreen, 2006; Bournaud et al., 2008; Shapiro et al., 2008) suggests that clumps form in gas rich spiral disks rather than during on-going mergers, although the latter scenario can not be completely ruled out (Taniguchi & Shioya, 2001; Overzier et al., 2008). Recent work by Bournaud et al. (2007) (hereafter B07) and Elmegreen et al. (2008) has demonstrated that internal disk fragmentation can effectively reproduce many of the observables of chain and clump clusters galaxies and that these different clumpy systems can have the same origin but observed at different inclinations. However, the models of B07

---

[26] This chapter has been published as Agertz et al. (2009) in Monthly Notices of the Royal Astronomical Society: Letters, Volume 397, Issue 1, pp. L64-L68.

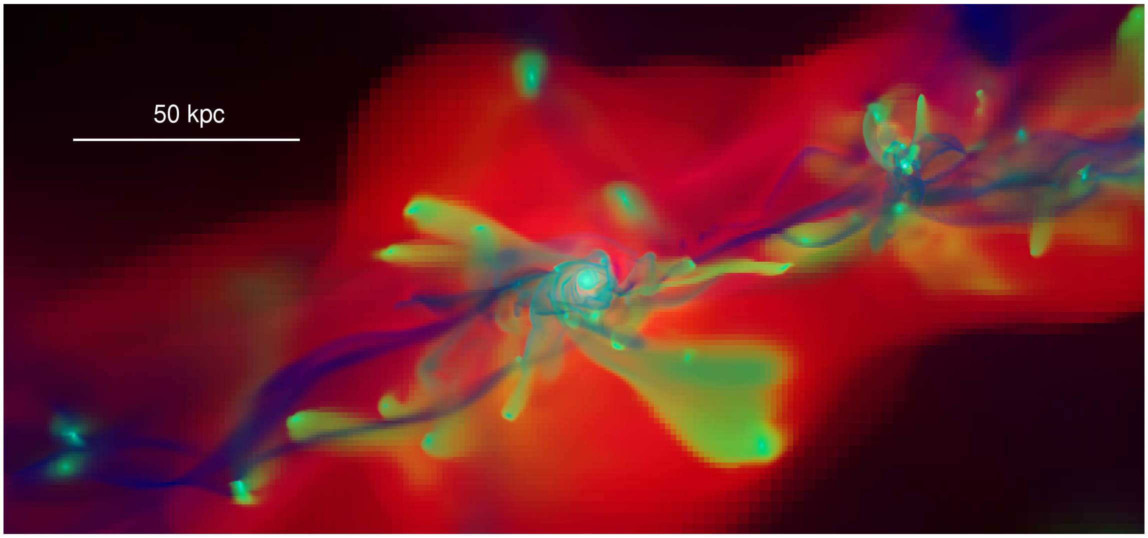

Figure 5.1: An *RGB*-image of the gas showing the disk and accretion region at $z \sim 3$. The image is constructed using *R*=temperature, *G*=metals and *B*=density. We can clearly distinguish the cold pristine gas streams in blue connecting directly onto the edge of the disk, the shock heated gas in red surrounding the disk and metal rich gas in green being stripped from smaller galaxies interacting with the halo and streams of gas. The disk and the interacting satellites stand out since they are cold, dense and metal rich.

still rely on idealized, pre-existing very massive gas disks, in order to reproduce the massive clumps and can not explain an ongoing, steady-state fragmentation scenario.

How galaxies acquire their baryons is an open question. The classic picture of galaxy formation within the cold dark matter (CDM) scenario assumes that the accreted gas is shock heated to the virial temperature, cools radiatively and rains down to form an inner star-forming rotating disk. Recent theoretical studies (Birnboim & Dekel, 2003; Kereš et al., 2005; Dekel & Birnboim, 2006; Ocvirk et al., 2008; Kereš et al., 2009; Brooks et al., 2009; Dekel et al., 2009) have demonstrated that accretion of fresh gas via cold infall can in fact be the dominant process for gas accretion for halo masses $M \lesssim 10^{11.6}\,M_{\odot}$. In these halos, the cooling time for gas of temperature $T \sim 10^4$ K is shorter than the timescale of gas compression and shocks are unable to develop. In halos above this mass, cold accretion persists as gas is supplied by cold streams penetrating through hot massive halos at $z \gtrsim 2$ (Ocvirk et al., 2008; Dekel et al., 2009) whilst the classical hot mode of gas accretion dominates at lower z. Because of insufficient spatial resolution, these studies could not follow the evolution of the accreting gas and how the cold streams connect to the central galaxies. The purpose of this *Letter* is to look in detail at the gas accretion and disk formation process using state-of-the-art numerical simulations.

5.2. Numerical Simulation

We use the adaptive mesh refinement (AMR) code RAMSES (Teyssier, 2002) to simulate the formation of a massive disk galaxy in a cosmological context including dark matter, gas and stars. The gas dynamics are calculated using a second-order unsplit Godunov method, while collisionless particles (including stars) are evolved using the Particle-Mesh technique. The modelling includes realistic recipes for star formation (Rasera & Teyssier, 2006), supernova feedback and enrichment (Dubois & Teyssier, 2008). Metals are advected as a passive scalar and are incorporated self-consistently in the cooling and heating routine, as in Agertz et al. (2009), and we adopt an initial metallicity of $Z = 10^{-3} Z_{\odot}$ in the high-resolution region. The refinement strategy is based on a quasi-Lagrangian approach, so that the number of particles per cell remains roughly constant, avoiding discreteness effects (e.g. Romeo et al., 2008). The computational domain is a 40 Mpc cube containing nested AMR grids of particles and gas cells down to a Lagrangian region containing dark matter particles of mass $m_{\rm p} = 2.2 \times 10^5\,M_{\odot}$. The effective resolution of our initial grid is therefore 2048^3. We then refine this base grid according to our refinement strategy, so that the maximum resolution is $\Delta x \sim 40$ pc in *physical units* at all times.

For our initial conditions we take the Via-Lactea II simulation (Diemand et al., 2008) which forms a Milky Way sized dark matter halo that accretes most of its mass ($M_{\rm vir} = 2 \times 10^{12}\, M_\odot$ at $z = 0$) by redshift $z = 2$. We evolved the entire simulation to $z = 0$ at a coarser resolution, here we report on the high redshift evolution to $z = 2$ at which point it hosts a disk that is massive enough to be compared to the observations in e.g. Bournaud et al. (2008) and Genzel et al. (2006). We use standard galaxy formation ingredients, with a star formation efficiency of 2% (as defined in Rasera & Teyssier, 2006), a star formation density threshold $n_{\rm H} = 4\,{\rm cm}^{-3}$ and a supernovae mass loading factor $f_w = 10$ (as defined in Dubois & Teyssier, 2008). In order to prevent artificial fragmentation, we use a pressure floor $P \simeq 3G\Delta x^2 \rho^2$, so that we satisfy the Truelove et al. (1997) criterion at all times.

5.3. Results

Fig. 5.1 shows a large scale view of the galactic disk at $z \sim 3$. At this time the dark matter halo has reached a mass of $M_{\rm vir} \sim 3.5 \times 10^{11}\, M_\odot$, while the total baryonic mass in the disk (disregarding the bulge) is $M_{\rm bar} \sim 2.4 \times 10^{10}\, M_\odot$ out of which 50% is gas, putting it in a regime where both cold flows and stable shocks can exist (Kereš et al., 2005; Dekel & Birnboim, 2006; Ocvirk et al., 2008). This striking image ties together many aspects present in modern theories of galaxy formation and highlights new complexities. Cold streams of gas originating in narrow dark matter filaments, effectively penetrate the halo and transport cold metal-poor gas right down to the proto-galactic disk to fuel the star forming region. A comparable amount of metal enriched material reaches the disk in a process that has previously been unresolved - material that is hydrodynamically stripped from accreting satellites, themselves small disky systems, through the interaction with the hot halo and frequent crossings of the cold streams.

Streams of cold gas flow into the halo on radial trajectories, eventually forming orderly rotational motion in an extended disk. This gas is in approximate pressure equilibrium with the hot halo that has a rotational velocity of $v_{\rm rot} \sim 30\,{\rm km\,s^{-1}}$ close to the virial radius, increasing smoothly to $v_{\rm rot} \sim 200\,{\rm km\,s^{-1}}$ to match the rotation at the edge of the disk ($r \sim 10\,{\rm kpc}$). The ram pressure is significant close to the disk, forcing the streams to curve around it. At early times, when the interaction region close to the disk is tenuous, streams can 'swing' past the proto-disk before being decelerated completely. At later times the turbulent accretion region carries significant mass and infalling cold gas quickly decelerates by plowing through it. We detect compression and radiative shocks that quickly dissipate since the cooling times are very short, resulting in a denser configuration for the cold gas. The global outcome of these interaction is a turbulent gas heavy disk prone to fragmentation. Fig. 5.2 shows a time sequence of the complicated and asymmetric gas flows around the gas disk at $z \sim 3$. The figure reveals that many of the large scale spiral arms at large radii are not waves, but material arms that can survive for an orbital time and that these arms are gravitationally unstable and can fragment into clumps. Gravitational instability has been used by Elmegreen et al. (1993) (hereafter E93) to explain the formation of massive clumps, as large as dwarf galaxies, in the tidal tails of merging galaxies. The typical mass of objects that form within the arms is $M_{\rm J} \simeq \sigma_{\rm eff}^4/G^2\Sigma$, where Σ is the surface density of gas within the arm and the effective mass-weighted 1D velocity dispersion is defined as $\sigma_{\rm eff}^2 = c_{\rm s}^2 + \sigma_{\rm 1D}^2$ where $c_{\rm s}$ is the local sound speed. Using the small region highlighted by the grey square in Fig. 5.2, we have measured $\Sigma = 60\, M_\odot\,{\rm pc}^{-2}$ and $\sigma_{\rm eff} \simeq 25\,{\rm km\,s^{-1}}$, giving $M_{\rm J} \simeq 2 \times 10^8\, M_\odot$ which agrees well with the mass of the forming clump. The internal dispersion velocity is roughly equal to the divergent motions across the curved gas filament. The typical velocity dispersion across a $\lambda \sim 1$ kpc patch of the filament will be of the order $\sigma \simeq \lambda v_{\rm orb}/\mathcal{R}_{\rm c} \simeq 20\ {\rm km\,s^{-1}}$, where the orbital velocity $v_{\rm orb} \sim 200\ {\rm km\,s^{-1}}$ and the curvature radius $\mathcal{R}_c$ of the filament equals the radius of the extended disk. This value agrees well with the dominating turbulent component of $\sigma_{\rm eff}$. As the interaction region grows in mass and develops a more symmetric disk-like morphology we also observe massive clump formation in fragmenting spiral waves at intermediate radii. The resulting galaxy is shown in Fig. 5.3 at $z \sim 2.7$, after many large clumps have formed through the above mechanisms. We detect 14 clumps with masses between $M_{\rm cl} \sim 5 \times 10^7$ and $10^9 M_\odot$, of which only the two smallest did not form *in situ* but were infalling satellites. The three most massive clumps have $M_{\rm cl} \sim 7 - 8 \times 10^8\, M_\odot$. In total $\sim 15\,\%$ of the baryons are in clumps. In the interaction region between the disk and the cold streams, the typical arm surface density and velocity dispersion can both be estimated

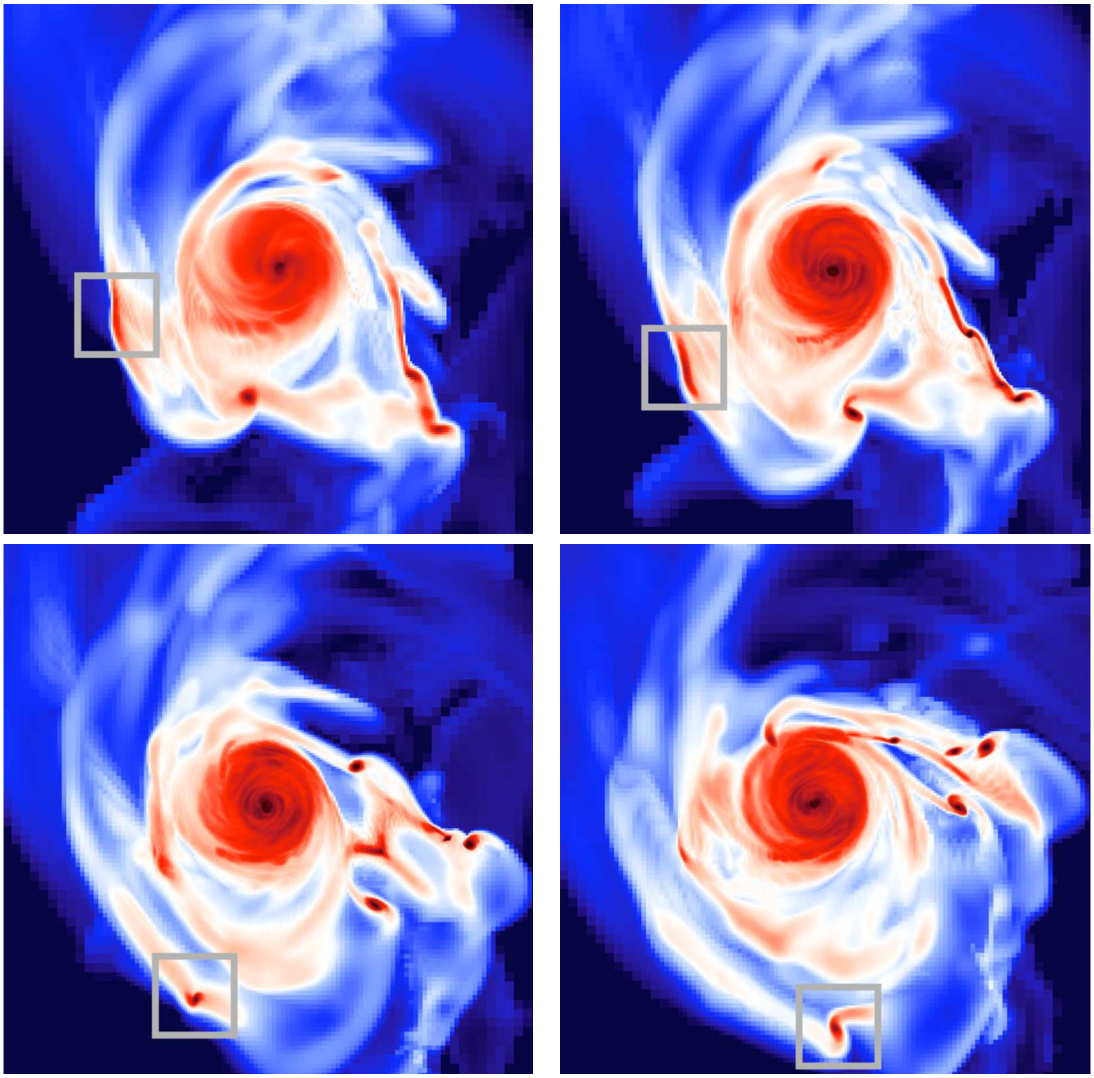

Figure 5.2: A time sequence spanning 40 Myr of the projected gas density at $z \sim 3$ in a 18×18 kpc^2 region. The box shows the formation of a $\sim 10^8\, M_\odot$ clump via gravitational instability.

using mass average quantities within cylindrical shells. At ($r \sim 8$ kpc) we measure $\langle\Sigma\rangle \simeq 20$ $\mathrm{M_\odot\ pc^{-2}}$ and $\langle\sigma\rangle \simeq 30$ km $\mathrm{s^{-1}}$, giving rise to clump masses as large as $M_\mathrm{J} \simeq 10^9\, M_\odot$. Even though we satisfy the Truelove criterium, convergence in the details of the clump properties can be influenced by numerical fragmentation and may require more cells per Jeans length. In addition, numerical diffusion from bulk flows can lead to an underestimation of the turbulent velocity dispersion. Quantifying this is beyond the scope of this paper. The detected clumps are located in the interaction region between the inner disk and the cold streams. In our case this region is not aligned with the initial galactic disk, giving rise to a misalignment of the clumps with respect to the inner galactic disk (see edge-on images in Fig. 5.3). Although we believe that this misalignment is not typical, it is an elegant explanation for the formation of 'bent' chains, such as the one reported in Bournaud et al. (2008). Indeed, Elmegreen & Elmegreen (2006) report that the typical chain galaxy has clumps mostly aligned in the midplane, while in some cases, clumps are seen above and below the midplane (outer and inner disk misaligned). In our case, the misalignment is due to a third cold stream that is perpendicular to the main filament seen in Fig. 5.1. In a similar scenario, this process has also been invoked to explain the formation of large polar rings (Macciò et al., 2006).

The simulated galaxy is sharing many properties with observed chain and clump cluster galaxies (Elmegreen

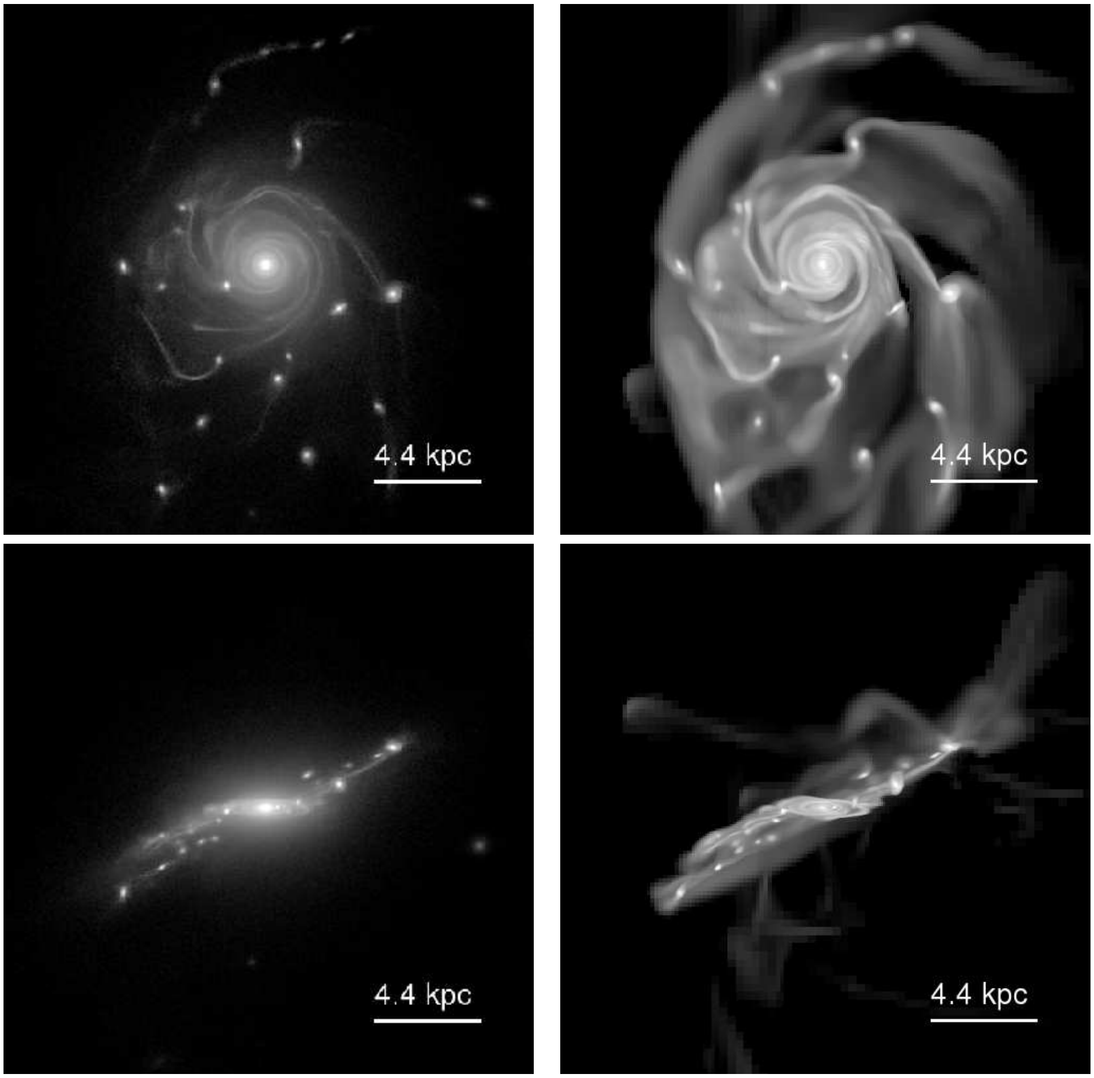

Figure 5.3: Density projection of the stars (left-hand panels) and gas (right-hand panels) at $z \sim 2.7$ illustrating the fragmentation process and the formation of large clumps of mass $\sim 10^7 - 10^9\,M_\odot$.

et al., 2007). Viewed edge-on, the misaligned disk morphology is clearly seen and the overall structure resembles a large chain-galaxy. Viewed face on the spiral-like structure has a similar morphology as clump clusters or clumpy spirals. Elmegreen & Elmegreen (2005) report that UDF clumpy galaxies at $z \sim 1.5 - 3$ have a stellar mass $\simeq 6 \times 10^{10}\,M_\odot$ and a radius ~ 10 kpc, in striking agreement with our simulated galaxy. Not only does the cosmological simulation reproduce the observed clumpy morphology and global rotation of these systems but we also find a realistic metallicity gradient and star formation rate of 20 $M_\odot \mathrm{yr}^{-1}$. The inner disk has on average solar metallicity, while that in the clump forming region is only $\sim 1/10\,Z_\odot$, due to the accretion of pristine gas in the cold streams mixing with stripped satellite gas. This has the important observational consequence that these massive clumps might be devoid of dust, making them easier to detect. To illustrate 'how disks acquire their baryons', we have plotted the mass accretion rate in different gas phases measured around our simulated galaxy at $z = 5, 3$ and 2 in Fig. 5.4. We define the phases as cold diffuse ($T < 2 \times 10^5$ K, $n < 0.05\,\mathrm{cm}^{-3}$), dense ($n > 0.05\,\mathrm{cm}^{-3}$), hot diffuse ($T > 2 \times 10^5$ K, $n < 0.05\,\mathrm{cm}^{-3}$) and stripped ($Z > 0.01\,Z_\odot$, $n < 0.05\,\mathrm{cm}^{-3}$). Indeed, at $z = 3$ and 5 the mass accretion rates in cold streams is very high ($\dot{M} \simeq 20\,M_\odot \mathrm{yr}^{-1}$). A significant amount of baryons are also accreted from stripped satellites, although quantifying this amount is difficult in Eulerian schemes since this metal rich material can mix with

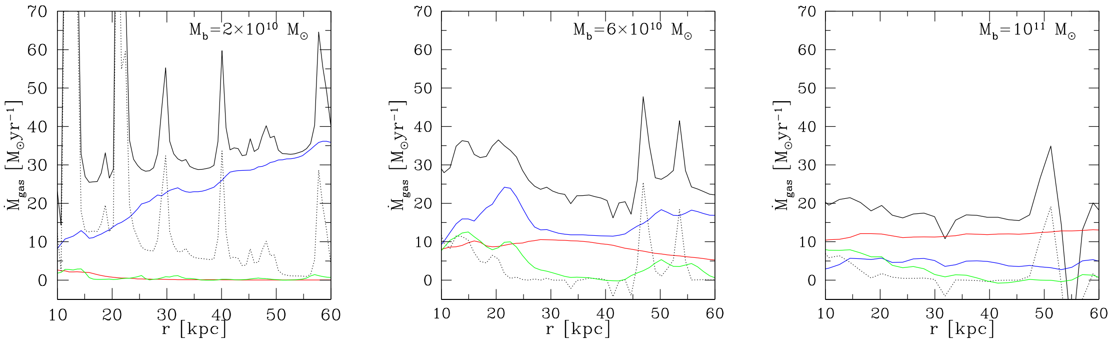

Figure 5.4: Mass accretion averaged within spherical shells at redshifts $z \sim 5, 3$ and 2. The radii are in physical kpc. The lines show the total mass flow (solid black) in each shell, cold diffuse (blue solid), hot diffuse (red solid), dense (dotted) and stripped gas (green) (see text for definition). We observe a decrease in the overall inflow of material and a change from cold to hot accretion over time.

the other gas phases that have never been part of the satellites. After $z \sim 2$, the hot mode of accretion dominates, making large clump formation at large radii only possible through galaxy mergers, c.f. Barnes & Hernquist (1992) and E93. At $z \sim 2$, the galaxy has a thin and extended spiral disk component. Although the gas velocity dispersion is still rather high in the disk, the Jeans mass in the spiral arms is on the order of $\sim 10^7\ M_\odot$, closer to the largest giant molecular clouds in present day spiral galaxies. The corresponding gas Q_g-parameter (Goldreich & Lynden-Bell, 1965a) is $Q_\mathrm{g} \simeq 1.5-2$ in the star forming region, indicating that the disk is marginally stable and the galaxy has reached a quiescent phase with no further large clump formation.

Fig. 5.5, shows the dark matter mass accretion rate in the simulated galaxy, as a function of time. At $z = 2$, the accretion rate is significantly lower than the average, explaining why the disk has reached this quiescent phase. A global analytical approach for understanding high-z disk fragmentation can be applied (Dekel et al., 2009), based on simple stability arguments and the disk fraction $\delta \equiv M_\mathrm{d}/M_\mathrm{tot}(R_\mathrm{d})$. Here M_d is the baryonic mass in the disk and $M_\mathrm{tot}(R_\mathrm{d})$ is the total mass within the disk radius R_d. A $\delta \sim 0.25-0.5$ should give rise to large clumps involving a few percent of the disk mass and $\delta \sim 0.3-0.35$ is predicted for a steady-state fragmentation from moderately clumpy streams. The disk in our simulation at $z = 5, 3$ and 2 has $\delta = 0.47, 0.33$ and 0.33 respectively which is in excellent agreement with the above prediction (see also Fig. 2 in Dekel et al. (2009)). Using only the gas in M_d renders a lower bound of $\delta = 0.17, 0.17$ and 0.1. We point out that the stellar fraction increases significantly towards lower redshifts and this 'hotter' component stabilizes the disk at $z \sim 2$.

5.4. Conclusions

We have followed the formation and accretion history of a Milky Way sized galaxy using state-of-the-art AMR techniques. Most of the baryons are in an orderly rotating disk by a redshift $z = 2$, but how they attain this equilibrium is very complex and the focus of this work. One of the most important points of this paper is that we can answer the question in detail of 'how galaxies get their baryons'. Extending recent work on the impact of cold streams on galaxy formation (Kereš et al., 2005; Dekel & Birnboim, 2006; Ocvirk et al., 2008; Dekel et al., 2009), we analyze for the first time how single phase narrow cold streams and ram pressure stripped debris assembles an extended turbulent rotating disk. Complex gas interactions takes place in an extended accretion region in which infalling gas is decelerated through compression/radiative shocks and from the pressure gradients arising from a hot halo component.

Prior to $z \sim 2$, the accretion rate of cold gaseous material onto the disk is the highest and we resolve the gravitational instabilities responsible for the formation of many very massive clumps within an extended ~ 10 kpc disk. This is about two times larger than the theoretical expectations of disk sizes at this epoch

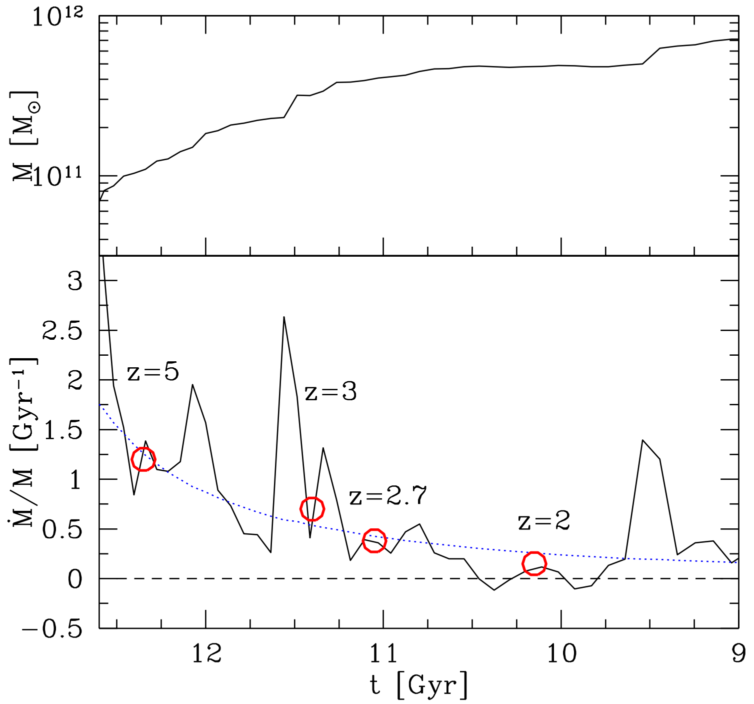

Figure 5.5: Top panel: dark matter accretion history [and its logarithmic derivative (bottom panel)] for the VL-2 halo using the HOP halo finder (Eisenstein & Hut, 1998). After a period of moderate mass increase, during several epochs (e.g. $z \sim 3.25$ and ~ 1.75) the halo mass and hence the gaseous mass dramatically increases. The red rings mark specific times discussed in the text. The dotted blue line shows the expected averaged gas accretion calculated from extended Press-Schechter (EPS) theory (Neistein et al., 2006).

(Mo et al., 1998). The observed morphology, star forming rate, global rotation and metallicity of the system is in good agreement with the observed clump-cluster and chain galaxies (Elmegreen & Elmegreen, 2006; Elmegreen et al., 2007; Bournaud et al., 2008). This scenario is an extension of the disk fragmentation scenario proposed by B07 and Elmegreen et al. (2008), although here studied more consistently within the current cosmological framework, and more specifically related to cosmological accretion. Therefore, clumpy galaxies should be most frequent at this epoch since massive clump formation stops during the remaining slow accretion phase and the disk evolves quiescently until $z = 0$ which will be reported on in a forthcoming paper.

Acknowledgments

The authors would like to thank Aaron Boley for valuable discussions on gravitational instability, Fredric Bournaud for discussions on clumpy galaxies and the referee Avishai Dekel for a careful review that improved the quality of the paper. Thanks to Jürg Diemand for making the initial conditions available and to Doug Potter for making it possible to run the simulations on the zBox supercomputers (http://www.zbox2.org).

6.

The Formation of Disk Galaxies in a ΛCDM Universe[27]

We study the formation of disk galaxies in a fully cosmological framework using adaptive mesh refinement simulations. We perform an extensive parameter study of the main sub-grid processes that control how gas is converted into stars and the coupled effect of supernovae feedback. We argue that previous attempts to form disk galaxies have been unsuccessful because of the universal adoption of strong feedback combined with high star formation efficiencies. Unless extreme amounts of energy are injected into the interstellar medium during supernovae events, these star formation parameters result in bulge dominated S0/Sa galaxies as star formation is too efficient at $z \sim 3$. We show that a low efficiency of star-formation more closely models the sub-parsec physical processes, especially at high redshift. We highlight the successful formation of extended disk galaxies with scale lengths $r_{\rm d} = 4-5\,{\rm kpc}$, flat rotation curves and bulge to disk ratios of B/D$\sim 1/4$. Not only do we resolve the formation of a Milky Way-like spiral galaxy, we also observe the secular evolution of the disk as it forms a pseudo-bulge. The disk properties agree well with observations and are compatible with the photometric and baryonic Tully-Fisher relations, the $\Sigma_{\rm SFR}-\Sigma_{\rm gas}$ (Kennicutt-Schmidt) relation and the observed angular momentum content of spiral galaxies. We conclude that underlying small-scale star formation physics plays a larger role than previously considered in simulations of galaxy formation.

6.1. Introduction

The prevailing picture of galaxy formation emerged more than 30 years ago (White & Rees, 1978; Fall & Efstathiou, 1980). Within the framework of the broadly accepted Λ Cold Dark Matter (ΛCDM) scenario (Komatsu et al., 2009), gravity assembles structures in a bottom-up fashion. Haloes of dark matter acquire angular momentum via tidal torques (Peebles, 1969; Fall & Efstathiou, 1980) from interacting structures, and as gas cools and condenses into their central parts, star forming galaxies form. A realistic angular momentum content can be accounted for if most of the angular momentum is retained in the assembly process. In this picture, the host halo is responsible for the final galaxy characteristics (e.g. Mo et al., 1998). While several aspects of the theory of galaxy formation are still being developed, e.g. the underlying physics of the missing satellite problem (Klypin et al., 1999; Moore et al., 1999) and the role of cold stream accretion (Kereš et al., 2005, 2009; Dekel et al., 2009), the model has proven successful for understanding global properties of galaxy assembly.

Given the complexity and non-linearity of the involved processes, computer simulations have become the ideal tool for studying the formation of structure. The formation of a late-type spiral galaxy, such as our own Milky Way, has been studied numerically in fully ΛCDM cosmological context by many authors (e.g. Abadi et al., 2003b; Sommer-Larsen et al., 2003; Governato et al., 2004; Robertson et al., 2004; Okamoto et al., 2005; Governato et al., 2007; Croft et al., 2009; Scannapieco et al., 2009; Piontek & Steinmetz, 2009b; Agertz et al., 2009). To date, no attempt has yielded a realistic candidate. The dominant reason for this is the so called "angular momentum problem" which leads to small, centrally concentrated disks dominated by large bulges (Navarro & Benz, 1991; Navarro & White, 1994). Merging substructures lose angular momentum to

---

[27] This chapter has been published as Agertz et al. (2011) in Monthly Notices of the Royal Astronomical Society.

the outer halo via dynamical friction, forcing the associated baryons to end up in the central parts of the protogalaxy as a spheroid rather than a disk. This poses a problem for the theoretical understanding of extended late-type galaxies. This might in part stem from numerical issues: the commonly used Smoothed Particle Hydrodynamics (SPH) (Gingold & Monaghan, 1977; Lucy, 1977) technique is known to incorrectly treat boundaries, hence poorly treating multiphase fluids (e.g. Agertz et al., 2007; Read et al., 2010). This can lead to artificial angular momentum transfer at the interface between cold disk and a hot halo (Okamoto et al., 2005).

Many proposed solutions exists to the angular momentum problem, all amounting to the same process: keep the gas from cooling and forming stars too efficiently in the merging dark matter satellites at high redshift. One natural source is the cosmological UV background, being responsible for reionization at $z \gtrsim 6$ which heats the gas, preventing it to cool efficiently into star forming dwarf galaxies (Thoul & Weinberg, 1996; Quinn et al., 1996; Gnedin, 2000; Hoeft et al., 2006). However, the impact on objects larger than $v_{\rm circ} \sim 10\,{\rm km\,s^{-1}}$ is unclear due to e.g. self-shielding and efficient collisional cooling (Dijkstra et al., 2004).

Gas in low mass haloes can also be blown out by supernova driven winds (Dekel & Silk, 1986; Efstathiou, 2000), hence lowering the resulting star formation efficiency (SFE), enriching the IGM in the process. Mac Low & Ferrara (1999) demonstrated that while dwarf galaxies of mass $10^6 - 10^9\,{\rm M_\odot}$ efficiently can expel metals in supernovae driven winds, virtually no mass is lost for systems of mass $\gtrsim 10^7\,{\rm M_\odot}$ (see also Dubois & Teyssier, 2008). The inefficiency in driving winds from dwarfs was also reported by Marcolini et al. (2006) who attributed this to the extended dark matter halo and efficient metal cooling. In this scenario, mass loss and IGM enrichment will occur due to tidal and ram-pressure stripping (e.g. Mori & Burkert, 2000). Phenomenological models of e.g. momentum driven winds have proven successful in reproducing the high-z IGM (Oppenheimer & Davé, 2006) but it is uncertain how it regulates star formation and in what manner the expelled gas is re-accreted at later times (Oppenheimer et al., 2010).

Various recipes of supernovae feedback have been developed for numerical simulations (e.g. Navarro & White, 1993; Kay et al., 2002; Scannapieco et al., 2006), and the methods have proven successful in removing low angular momentum material from central parts of galaxies (e.g. Sommer-Larsen et al., 2003; Okamoto et al., 2005; Governato et al., 2007), yielding more extended galaxies in comparison to models without feedback. However, it is unclear to what extent this way of reducing star formation can account for disk dominated spiral galaxies like the Milky Way. Recently Scannapieco et al. (2009) demonstrated numerically, in a fully cosmological setting, how a set of 8 Milky Way sized haloes failed to form significant disks. While half of the sample were early type galaxies resulting from late time mergers, the other half of the sample had less than 20 per cent of their stellar mass in disks. This can be a result of the inability of the adopted feedback to remove or redistribute low angular momentum material, but is also a strong indication that something else might regulate star formation at high redshift. On the same topic, Sawala et al. (2010) argues that modern simulations of dwarf galaxy formation (Valcke et al., 2008; Stinson et al., 2009; Governato et al., 2010) all yield much larger stellar masses than expected from observations as well as gas-to-star conversion efficiencies almost an order of magnitude too large. Dutton & van den Bosch (2009) found that, for SNe feedback to yield realistic galaxies, it must be very efficient, converting 25 per cent of the SN energy into outflows. If too strong feedback is employed, the disks can be destroyed by internal processes as too much material is ejected into the halo, preventing efficient disk reformation from cold gas, and possibly violating the upper bounds of halo gas found in X-ray surveys (see Bregman (2007) and references within). In light of these studies, it is unclear if supernovae feedback is the sole agent in regulating star formation. Note that SNe explosions can regulate star formation in galaxies without expelling gas, being a driver of galactic turbulence (Mac Low & Klessen, 2004).

Fundamentally, star formation is regulated by the availability of H_2. The observed Kennicutt-Schmidt (from now on K-S) relation (Kennicutt, 1998), that relates $\Sigma_{\rm SFR}$ to $\Sigma_{\rm gas}$, varies strongly among individual spiral galaxies and can not be fit with a single power law (Bigiel et al., 2008). $\Sigma_{\rm SFR}$ behaves very differently for $\Sigma_{\rm gas}$ greater or smaller than $\approx 9\,{\rm M_\odot\,pc^{-2}}$, marking the transition from atomic to fully molecular star forming gas (Leroy et al., 2008), and is dependent on gas metallicity, dust content, turbulence, small scale clumpiness and local dissociating UV field (McKee & Ostriker, 2007). The inclusion of these processes and its impact on global star formation in disks has recently been studied both numerically (Robertson & Kravtsov, 2008; Gnedin et al., 2009; Pelupessy & Papadopoulos, 2009) as well as analytically (e.g. Krumholz et al., 2009). A natural outcome of this treatment is an order of magnitude lower amplitude of the K-S relation at high redshifts ($z \sim 3$) (Gnedin & Kravtsov, 2010b). This agrees well with the observation of damped Lyα systems (DLA: Wolfe & Chen, 2006) as well as Lyman Break Galaxies (Rafelski et al., 2009). This indicates that star formation can be made inefficient at high redshift, leaving gas for late time star formation in a disk like environment, but not necessarily by expelling gas in supernova driven winds. In addition, Murray et al. (2010) argues that the disruption time-scale of giant molecular clouds (GMCs) due to jets, HII gas pressure, and radiation pressure also serves to regulate the SFE in galaxies. The disruption occurs well before the most massive stars exit the main sequence, meaning that supernovae in principle have little effect on GMC lifetimes.

In this paper we investigate to what extent supernovae feedback and the underlying small scale star forming physics can affect the formation and evolution of realistic spiral galaxies in a fully cosmological setting. The former effect is studied via well tested numerical implementations of SNII, SNIa feedback coupled to metal enrichment, as well as stellar mass loss. The latter influence is achieved by considering different normalizations of the Schmidt-law star formation efficiency. We conduct a comprehensive analysis of the resulting $z = 0$ disks and compare them to observational relations.

The paper is organized as follows. In Sect. 6.2, we describe the numerical method used in this work, including the adopted feedback and star formation prescriptions. In Sect. 6.3, we present the cosmological initial conditions and discuss the free parameters of this work. Sect. 6.4 outlines the disk analysis and summarizes the final properties of the simulation suite. In Sect. 6.5 and Sect. 6.6, we present a detailed analysis of the impact of small-scale SFE and supernova feedback respectively. In Sect. 6.7 we compare our simulations to modern observations. Finally, Sect. 6.8 summarizes and discusses our conclusions.

6.2. Numerical Framework

We use the Adaptive Mesh Refinement (AMR) code RAMSES (Teyssier, 2002) to simulate the formation of a massive disk galaxy in a cosmological context including dark matter, gas and stars. The gas dynamics is calculated using a second-order unsplit Godunov method, while collisionless particles (including stars) are evolved using the particle-mesh technique. The equation of state of the gas is that of a perfect mono-atomic gas with an adiabatic index $\gamma = 5/3$. Self-gravity of the gas is calculated by solving the Poisson equation using the multi-grid method (Brandt, 1977) on the coarse grid and by the conjugate gradient method on finer ones. The modeling includes realistic recipes for star formation (Rasera & Teyssier, 2006), supernova feedback and enrichment (Dubois & Teyssier, 2008). Details on these implementations are given below. Metals are advected as a passive scalar and are incorporated self-consistently in the cooling and heating routine. The code adopts the cooling function of Sutherland & Dopita (1993) for cooling at temperatures $10^4 - 10^{8.5}$ K. We extend cooling down to 300 K using rates form Rosen & Bregman (1995). Gas metallicity is also accounted for in the cooling routines. A UV background is considered using the prescription of Haardt & Madau (1996). In order to model a sub-grid gaseous equation of state, hence avoiding artificial gas fragmentation, the gas is given a polytropic equation of state

$$T = T_0 \left(\frac{\rho}{\rho_0}\right)^{\gamma_0 - 1}, \tag{6.1}$$

for densities large than ρ_0. Throughout this paper we adopt $T_0 = 1000\,\mathrm{K}$ and $\gamma_0 = 2.0$. In this work, the polytrope density is set equal to the star formation threshold n_0. Following Agertz et al. (2009), we adopt an initial metallicity of $Z = 10^{-4} Z_\odot$ in the high-resolution region. This also serves as a flag for allowed regions of refinement. The refinement strategy is based on a Quasi-Lagrangian approach, so that the number of particles per cell remains roughly constant, avoiding discreteness effects (e.g. Romeo et al., 2008).

Star Formation — To model the conversion of gas into stars we adopt a Schmidt-law (Schmidt, 1959) of the form

$$\dot{\rho}_g = -\epsilon_{\rm ff} \frac{\rho_{\rm g}}{t_{\rm ff}} \text{ for } \rho > \rho_0, \tag{6.2}$$

where $\rho_{\rm gas}$ is the gas density, $t_{\rm ff} = \sqrt{3\pi/32G\rho}$ is the local free-fall time, $\epsilon_{\rm ff}$ is the star formation efficiency per free-fall time and ρ_0 is the threshold for star formation. As soon as a cell is eligible for star formation, particles are spawned using a Poisson process where the stellar mass, m_*, is chosen to be a multiple of

$\rho_0 \Delta x^3$. Each formed star particles is treated as one stellar population with an associated initial mass function (IMF). This is a relevant approximation as the start particle masses are orders of magnitudes larger than the average stellar masses. We also ensure that no more than 90 per cent of the gas in a cell is depleted by the star formation process.

The ρ_0 and $\epsilon_{\rm ff}$ parameters in Eq. 6.2 are, in addition to being unconstrained physical parameters, resolution and hence scale dependent. There are in a sense two regimes of star formation in global simulations of disk galaxies:

1 – *The ISM is resolved:* At parsec scale resolution, star formation occurs in their natural sites i.e. massive clouds such as GMCs. Modern estimates of star formation efficiencies by Krumholz & Tan (2007) points towards values of $\epsilon_{\rm ff} = 1 - 2$ per cent at densities of $n \sim 10^2 - 10^5\,{\rm cm}^{-3}$. To only allow for star formation to take place in the actual physical star formation sites, hence tracing the formation of $\rm H_2$, allows for more accurate predictions of e.g. the Kennicutt-Schmidt star formation relation with less of a requirement to tune numerical parameters (see e.g. Gnedin et al., 2009). This treatment leads to $\rho_{\rm g} \rightarrow \rho_{\rm H_2}$ in Eq. 6.2, which is equivalent to $\epsilon_{\rm ff}$ being dependent on the environment, due in part to the local $\rm H_2$ fraction. On galactic scales, this means that the scale height of all ISM components are resolved using at least 10 resolution elements (Romeo, 1994). If this is not satisfied the true disk stability will not be modeled accurately. This treatment is the goal of most simulations, but is due to the computational load beyond the capabilities of modern simulations attempting to study the assembly and evolution of large spiral galaxies to $z = 0$. Isolated simulations of large spiral galaxies in a non-cosmological setting have successfully reached this resolution (Agertz et al., 2009; Tasker & Tan, 2009), albeit with simplified physics. As the star formation sites become resolved, new physics becomes important e.g. radiative feedback in order to accurately treat the life-times of GMC structures (Murray et al., 2010).

2 – *The ISM is under-resolved:* To radially resolve a Milky Way like galactic disk, i.e. sampling the scale radius with at least 10 resolution elements, a force and hydro resolution of a few $100\ {\rm pc}$ is necessary. At this resolution the scale height is captured with more or less one resolution element. The true disk stability can be affected as both the density and velocity structure (gas and stellar dispersion) are influenced numerically. This still allows for the disk to have the correct global properties such as gas and stellar mass compositions, thin and thick disk, and even to develop realistic spiral structure. In this case a statistical star formation recipe based on the local gas density and free fall time is well motivated both theoretically and observationally.

As we will describe in Sect. 6.3, we are targeting the latter regime of subgrid star formation and will investigate some of the numerical caveats related to it. At resolutions of several 100 pc, the $\epsilon_{\rm ff}$ parameter absorbs the small scale physics regulating star formation, allowing for a qualitative influence on galaxy formation. We note that there are alternative formulation of gas to star conversion laws in the literature (see e.g. Leroy et al. (2008) for a comprehensive summary). However, as they are all designed to fit an observed relation, and all include a normalization constant similar to $\epsilon_{\rm ff}$, we believe Eq. 6.2 is a representative choice for this study.

Supernovae and Stellar Feedback — The standard recipe for supernova feedback in RAMSES involves only Type II supernovae events (SNII). We have also implemented additional treatment of Type Ia events (SNIa) as well as mass loss via stellar winds. Including all of these effects is important as a single stellar population can return up to $30 - 40$ per cent of its mass to the ISM during its lifetime. The implementations are as follows:

□ *Type II* — Type II SN events are relevant for stellar masses of $8 - 40\ {\rm M}_\odot$ which represents ~ 10 per cent of the mass of a stellar population, regardless of IMF. We assume that 10 Myr after a star particle is formed, 10 per cent of the star particle's mass is injected into the nearest gas cell together with a total energy of $E_{\rm SNII} = 10^{51}(m_{\rm ejecta}/10\ {\rm M}_\odot)\,{\rm erg}$ in thermal energy. At low resolution, this energy would quickly radiate away (Katz, 1992) in the dense gas, without allowing for an adiabatic expansion of the supernova blast-wave (McKee & Ostriker, 1977a). To remedy this, we turn off cooling in cells containing young stars to allow for the blast-wave to grow and be resolved by few cells, hence converting thermal energy into $P{\rm d}V$ work (see e.g. Gerritsen, 1997). In detail, for every star formation event, the inverse of the birth time, $1/t_{\rm bt}$, is stored in the computational grid, overwriting any previous value. This passive scalar field is advected with the hydro flow (the conserved quantity is $\rho_{\mathrm{gas}}/t_{\mathrm{bt}}$). Gas cooling at a simulation time t_{sim} is only allowed if $t_{\mathrm{sim}} - t_{\mathrm{bt}} > \Delta t_{\mathrm{off}}$, where Δt_{off} is the cooling shut-off time-scale. Calculations of relevant time-scales and numerical tests using the SPH formalism was carried out by Stinson et al. (2006). Relevant time-scales are on the order of tens of Myr and we adopt $\Delta t_{\mathrm{off}} = 50\,\mathrm{Myr}$.

☐ *Type Ia* — We treat SNIa and stellar mass loss using the prescription outlined in Raiteri et al. (1996). The assumed IMF is the parametrization by Kroupa et al. (1993), which for a star particle of mass m_* reads

$$\Phi(M) = m_* A \begin{cases} 2^{0.9} M^{-1.3} & \text{for } 0.08 \leq M < 0.5\ \mathrm{M}_\odot \\ M^{-2.2} & \text{for } 0.5 \leq M < 1\ \mathrm{M}_\odot \\ M^{-2.7} & \text{for } M \geq 1\ \mathrm{M}_\odot, \end{cases}$$

where M is here the stellar mass in units of $\mathrm{M}_\odot$ and the normalization constant $A \approx 0.3029$. The adopted lower and upper limits are $0.08\,\mathrm{M}_\odot$ and $100\,\mathrm{M}_\odot$ respectively. At each simulation time step, we calculate the mass fraction of each star particle ending its H and He burning phase, i.e. leaving the main sequence, using the fit

$$\log t_* = a_0(Z) + a_1(Z)\log M + a_2(Z)(\log M)^2, \tag{6.3}$$

where t_* is the lifetime of the star and Z the metallicity. The adopted coefficients and references to the original data can be found in Raiteri et al. (1996). Progenitors of SNIa are carbon plus oxygen white dwarfs that accrete mass from binary companions. Stellar evolution theory predict the binary masses to be in the range of $\sim 3 - 16\,\mathrm{M}_\odot$. The number of SNIa events within a star particle, at a given simulation time with an associated timestep Δt, is

$$N_{\mathrm{SNIa}} = \int_{m_t}^{m_{t+\Delta t}} \hat{\Phi}(M_2)\mathrm{d}M_2, \tag{6.4}$$

where m_t and $m_{t+\Delta t}$ is the mass interval of stars ending their life during the computational timestep. $\hat{\Phi}(M_2)$ is the IMF of the secondary star, i.e.

$$\hat{\Phi}(M_2) = m_* A' \int_{M_{\mathrm{inf}}}^{M_{\mathrm{sup}}} \left(\frac{M_2}{M_{\mathrm{B}}}\right)^2 M_{\mathrm{B}}^{-2.7}\mathrm{d}M_{\mathrm{B}}, \tag{6.5}$$

where M_{B} is the mass of the binary, $M_{\mathrm{inf}} = \max(2M_2, 3\,\mathrm{M}_\odot)$ and $M_{\mathrm{sup}} = M_2 + 8\,\mathrm{M}_\odot$. The constant $A' = 0.16A$, and is a calibrated value for SNIa events in our Galaxy (van den Bergh & McClure, 1994). Each explosion is assumed to release 10^{51} erg (released as thermal energy in the nearest gas cell) and $0.76\,\mathrm{M}_\odot$ of metal enriched material ($0.13\,\mathrm{M}_\odot$ of ^{16}O and $0.63\,\mathrm{M}_\odot$ of ^{56}Fe) (Thielemann et al., 1986).

☐ *Stellar Mass Loss* — For each time step Δt, and star particle, we calculate the average stellar mass, $\langle M \rangle$, exiting the main sequence using Eq. 6.3. The mass loss during the Δt time-span is calculated using the best fit initial-final mass relation of Kalirai et al. (2008):

$$M_{\mathrm{wind}} = 0.891 - 0.394/\langle M \rangle. \tag{6.6}$$

At each time interval Δt, the total mass loss in winds is

$$M_{\mathrm{tot,wind}} = f_m m_* M_{\mathrm{wind}}, \tag{6.7}$$

where

$$f_m = \int_{m_t}^{m_{t+\Delta t}} \Phi(M)\mathrm{d}M. \tag{6.8}$$

The lost stellar mass enters the gaseous mass in the nearest cell and the gas metallicity is updated consistently with the star particle's metallicity.

6.3. Initial Conditions and Simulation Suite

The initial conditions used in this work are a subset the Silver River simulation suite (Potter et al. in preparation), aimed to study the pure dark matter assembly history of a Milky Ways size halo in much greater detail than here. We adopt a WMAP5 (Komatsu et al., 2009) compatible cosmology, i.e. a ΛCDM Universe with $\Omega_\Lambda = 0.73$, $\Omega_\mathrm{m} = 0.27$, $\Omega_\mathrm{b} = 0.045$, $\sigma_8 = 0.8$ and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

The upcoming work by Potter et al. will present the details concerning the initial condition generation. Briefly, a pure dark matter simulation was performed using a simulation cube of size $L_\mathrm{box} = 179\,\mathrm{Mpc}$. At $z = 0$, a halo of mass $M_{200,\mathrm{c}} \approx 9.7 \times 10^{11}\ \mathrm{M_\odot}$ was selected for re-simulation at high resolution, and traced back to the initial conditions at $z = 133$. $M_{200,\mathrm{c}}$ is the virial mass of the halo, defined as the mass enclosed in a sphere with mean density 200 times the critical value. The corresponding virial radius is $r_{200,\mathrm{c}} = 205\ \mathrm{kpc}$. By using a definition based on 200 times the background density we obtain $M_{200,\mathrm{bg}} = 1.25 \times 10^{12}\ \mathrm{M_\odot}$ and $r_{200,\mathrm{bg}} = 340\,\mathrm{kpc}$. When baryons are included in the simulations, the final *total* halo mass remains roughly the same.

The halo has a quiet merger history, i.e. it undergoes no major merger after $z = 1$, which favors the formation of a late type galaxy. A nested hierarchy of initial conditions for the dark matter and baryons was generated using the `grafic++`[28] code, where we allow for the high resolution particles to extended to 3 virial radii from the center of the halo at $z = 0$. This avoids mixing of different mass dark matter particles in the inner parts of the domain. In this work, we focus on two sets of resolutions from the Silver River suite, referred to as SR5 and SR6. The simulations are identical apart form the number of dark matter particles, and hence particle mass, as well as maximal AMR refinement. In SR6 the dark matter particle mass is $m_\mathrm{DM} = 2.5 \times 10^6\ \mathrm{M_\odot}$ and in SR5 $m_\mathrm{DM} = 3.2 \times 10^5\ \mathrm{M_\odot}$. The mesh is refined if a cell contains more than 8 dark matter particles, and similar criterion is employed for the baryonic component. At the maximum level of refinement, the simulations reach a physical resolution of $\Delta x = 170\ \mathrm{pc}$ and $\Delta x = 340\ \mathrm{pc}$ in SR5 and SR6 respectively.

The Free Parameters — The goal of this work is to study how the characteristics of disk galaxies change when standard numerical parameters governing star formation are modified. Following the discussion in Sect. 6.1, we consider star formation regulation in two different ways: small scale ($\sim 100\,\mathrm{pc}$) physics such as H_2 abundance, GMC turbulence, metallicity, radiative effects etc., or via energy injection from supernovae explosions leading to gas expulsion in galactic winds. The first mechanism is modeled by varying the Schmidt-law (Eq. 6.2) SFE, ϵ_ff, which acts on a cell by cell basis. The latter is studied by increasing the injected SNII energy, E_SNII. In addition we study the impact of star formation threshold n_0, but in less detail.

The traditional way of treating star formation in simulations of galaxy formation (e.g. Governato et al., 2007; Piontek & Steinmetz, 2009b) is to tune the SFE parameter using an isolated disk model to match the observed K-S relation, most commonly the fitting formula given by Kennicutt (1998) of $z = 0$ galaxies. In addition, the recipe for energy injection via supernovae and its efficiency is tuned simultaneously. These parameters are then used in fully cosmological simulation of galaxy formation. This scheme assumes that supernova explosions are the main sources of star formation regulation at high redshift. As argued in Sect. 6.1, the numerically assumed constant efficiency is strongly redshift dependent and a $z = 0$ tuning is likely to over-predict star formation in more metal poor environment at higher redshift. We treat ϵ_ff as a free, but constant, parameter and adopt $\epsilon_\mathrm{ff} = 1, 2$ or 5 per cent in the fully cosmological context. These values are in agreement with GMC estimates from Krumholz & Tan (2007). As we will demonstrate below, lower values than what traditionally is adopted is preferred in order to form late type galaxies. Note that $\epsilon_\mathrm{ff} \approx 2c_*$, where c_* is the efficiency parameter used in e.g. Governato et al. (2007) and Scannapieco et al. (2009) (defined via $t_\mathrm{dyn} = 1/\sqrt{4\pi G\rho}$). Values of $c_* = 0.05 - 0.1$ are commonly employed i.e. a few times, up to an order of magnitude larger than what we consider here.

The standard SNII feedback described in Sect. 6.2 is the baseline feedback in all of our simulations. In a subset of simulations we add the additional stellar mass loss and SNIa treatment. The high efficiency simulation ($\epsilon_\mathrm{ff} = 5$ per cent) is used as a template for the impact on the injected feedback SNII energy, which we set to $E_\mathrm{SNII} = 1, 2$ and $5 \times 10^{51}\,\mathrm{erg}$. Using energies that are several times larger than the canonical $10^{51}\,\mathrm{erg}$ might be perceived as unrealistic, but we believe it is illustrative to study the extreme cases of this

---

[28] `http://grafic.sourceforge.net/`

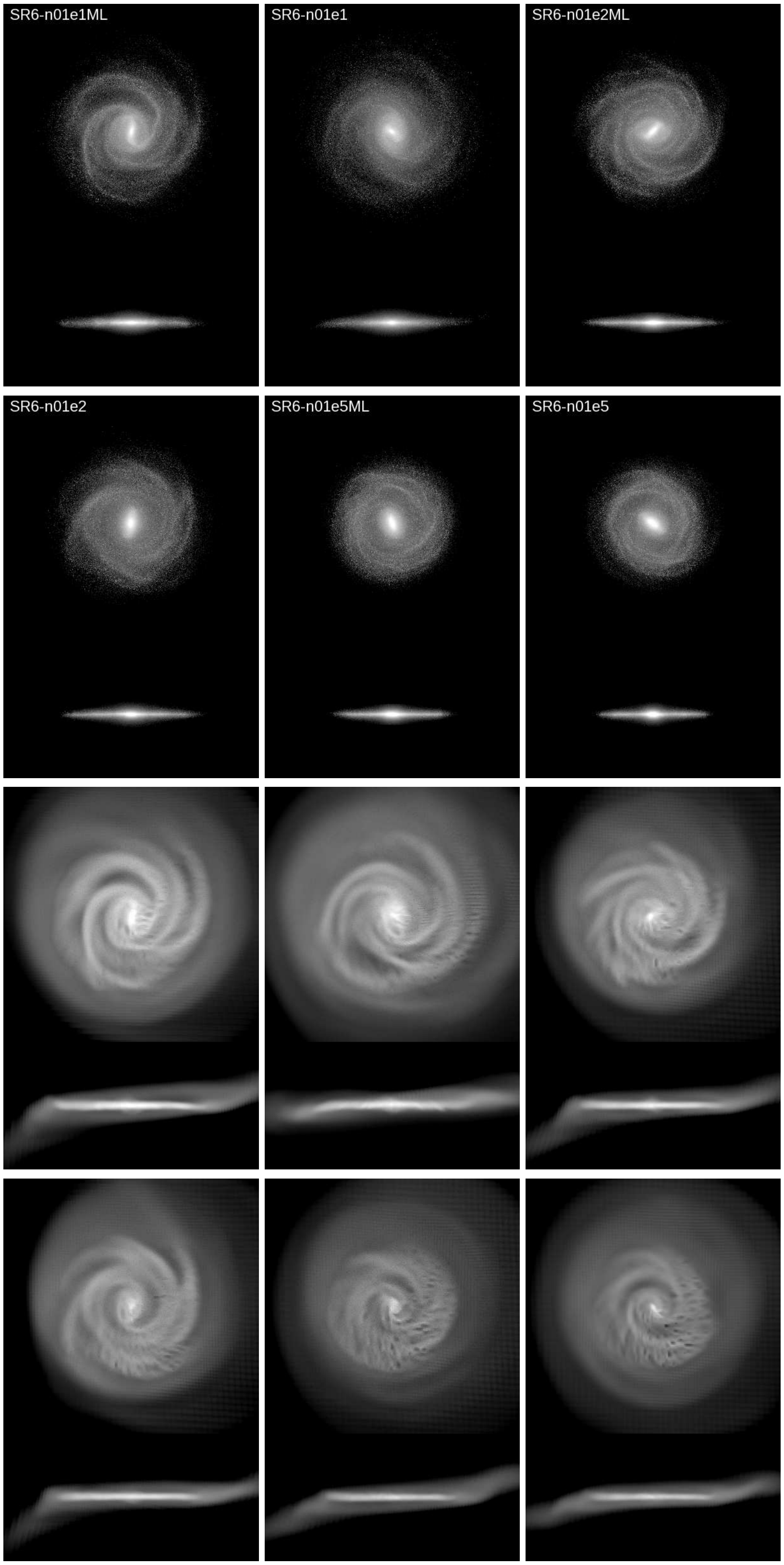

Figure 6.1: Projected face-on and edge-on surface density maps of the stars (top six panels) and gas (bottom six panels) of the $z = 0$ disks, where each panel is 60 kpc across. As the SFE is lowered and mass loss is employed, spiral structure becomes more pronounced due to a less massive bulge. The Hubble type of the disk changes from an early type (S0 or Sa) disk, to a late type spiral galaxy (Sb or Sbc) as we decrease $\epsilon_{\rm ff}$.

Table 6.1: Summary of the numerical parameters. The simulations use a maximum physical cell resolution of $\Delta x = 340\,\mathrm{pc}$ (SR6) or $\Delta x = 170\,\mathrm{pc}$ (SR5), and the high resolution region is occupied with dark matter particles of mass $m_{\mathrm{DM}} = 2.5 \times 10^6\,\mathrm{M}_\odot$ (SR6) or $m_{\mathrm{DM}} = 3.25 \times 10^5\,\mathrm{M}_\odot$ (SR5). All simulations use delayed cooling in regions of young stars, unless specified. When SNII feedback is used $E_{\mathrm{SNII}} = 10^{51}\,\mathrm{erg}$, unless other values are indicated.

Run	ϵ_{ff}	Feedback	Star formation threshold, n_0
SR6-n01e1	$1\,\%$	SNII	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e2	$2\,\%$	SNII	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e5	$5\,\%$	SNII	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e1ML	$1\,\%$	SNII, SNIa, mass loss	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e2ML	$2\,\%$	SNII, SNIa, mass loss	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e5ML	$5\,\%$	SNII, SNIa, mass loss	$0.1\,\mathrm{cm}^{-3}$
SR6-n1e1	$1\,\%$	SNII	$1\,\mathrm{cm}^{-3}$
SR6-n1e2	$2\,\%$	SNII	$1\,\mathrm{cm}^{-3}$
SR6-n1e5	$5\,\%$	SNII	$1\,\mathrm{cm}^{-3}$
SR6-n1e1ML	$1\,\%$	SNII, SNIa, mass loss	$1\,\mathrm{cm}^{-3}$
SR6-n1e2ML	$2\,\%$	SNII, SNIa, mass loss	$1\,\mathrm{cm}^{-3}$
SR6-n01e1NFB	$1\,\%$	No feedback, $E_{\mathrm{SNII}} = 0$	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e5NFB	$5\,\%$	No feedback, $E_{\mathrm{SNII}} = 0$	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e5NFBmet	$1\,\%$	No feedback but metal enrichment	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e5SN2	$5\,\%$	SNII, $E_{\mathrm{SNII}} = 2 \times 10^{51}\,\mathrm{erg}$	$0.1\,\mathrm{cm}^{-3}$
SR6-n01e5SN5	$5\,\%$	SNII, $E_{\mathrm{SNII}} = 5 \times 10^{51}\,\mathrm{erg}$	$0.1\,\mathrm{cm}^{-3}$
SR5-n1e1ML	$1\,\%$	SNII, SNIa, mass loss	$1\,\mathrm{cm}^{-3}$

type of feedback. In addition, the amount of SNII energy dissipated in cooling, after the shut-off time has passed, is complicated to measure. As a control set we also run the simulations without feedback, both with and without metal enrichment.

The philosophy of the star formation threshold is as follows. In reality stars form in molecular clouds of average densities of $n > 10^2\,\mathrm{cm}^{-3}$. Imposing a threshold of this magnitude would require a resolution on the order of parsecs to resolve the formation of the star forming clouds, something that is beyond the scope of fully cosmological hydro+N-body simulations today (but see (Gnedin et al., 2009)). We adopt $n_0 = 0.1$ and $1\,\mathrm{cm}^{-3}$ for each setting of ϵ_{ff}, but the appropriate choice is fundamentally tied to resolution and can lead to spurious results. The ISM has been shown to be represented by a lognormal density probability distribution function (PDF) (e.g. Kravtsov, 2003; Wada & Norman, 2007), or even a superposition of several log normally distributed ISM phases (Robertson & Kravtsov, 2008). The amount of gas eligible for star formation is represented by the high density part of the PDF which in turn is a function of total disk gas mass and turbulence. A density threshold should be picked to allow for the high-density star forming part of the PDF to be well-resolved or at least contains, given an adopted numerical resolution, a converged amount of star forming mass. If not, then chosen threshold will affect the numerical efficiency for global star formation. We will demonstrate this effect below.

In summary, the varied constants of interest is here the star formation threshold (n_0), the star formation efficiency per free fall time (ϵ_{ff}) and the form of supernova feedback and injected energy (E_{SNII}). Our main focus is the impact of these parameters at the SR6 level of resolution, and we present a brief resolution study in Appendix 6.9. We summarize our complete test suite in Table 6.1.

Table 6.2: Summary of disk characteristics at $z = 0$. The mass of the components are obtained by fitting the stellar surface density (see text), and are in units of $10^{10}\,\mathrm{M}_\odot$. Note that we consider *all* gas phases for the gas mass and all stars for the stellar masses. (1) Fitted scale length of stellar disk. Large uncertainties exist for $r_\mathrm{d} > 10$ kpc as the stellar disks are small and feature almost flat stellar surface density profiles. (2) $f_\mathrm{gas} = M_\mathrm{disk,g}/(M_\mathrm{disk,tot} + M_\mathrm{bulge})$. (3) Total measured specific angular momentum of the baryons in the disk and bulge in units of $\mathrm{km\,s^{-1}\,kpc}$.

Run	$M_\mathrm{disk,s}$	$M_\mathrm{disk,g}$	$M_\mathrm{bulge,s}$	r_d [kpc] (1)	f_gas (2)	B/D	B/T	j_bar (3)
SR6-n01e1	8.6	1.6	2.0	3.8	0.13	0.23	0.19	1920
SR6-n01e2	7.4	1.3	4.6	7.6	0.10	0.62	0.38	1655
SR6-n01e5	5.6	0.72	7.0	~ 15.0	0.05	1.25	0.56	1305
SR6-n01e1ML	8.0	2.3	2.2	5.0	0.18	0.27	0.21	1960
SR6-n01e2ML	8.1	1.6	3.8	5.0	0.12	0.47	0.32	1718
SR6-n01e5ML	5.5	0.93	7.2	~ 15.0	0.07	1.30	0.57	1464
SR6-n1e1	6.6	3.3	2.9	2.7	0.26	0.44	0.31	1594
SR6-n1e2	6.4	2.4	4.3	2.5	0.18	0.67	0.40	1804
SR6-n1e5	6.0	2.1	5.2	2.7	0.16	0.87	0.46	1643
SR6-n1e1ML	6.5	3.6	2.7	2.7	0.28	0.42	0.29	1618
SR6-n1e2ML	6.3	2.9	4.3	2.7	0.21	0.68	0.41	1281
SR6-n01e1NFB	7.8	1.6	3.1	4.0	0.13	0.40	0.28	1938
SR6-n01e5NFB	5.8	0.62	6.5	~ 15.0	0.05	1.12	0.53	1394
SR6-n01e5NFBmet	5.6	0.56	6.6	10.0	0.05	1.18	0.54	1430
SR6-n01e5SN2	5.6	0.55	6.5	4.5	0.05	1.16	0.54	1266
SR6-n01e5SN5	6.6	0.81	2.3	2.8	0.09	0.35	0.26	1387
SR5-n1e1ML	9.0	2.0	2.2	2.8	0.18	0.24	0.20	

6.4. The Disks

In this work, we focus primarily on the disk properties in the SR6 simulations at $z = 0$. Details of the satellite galaxies and halo properties will be considered in a future work. Fig. 6.1 and Fig. 6.2 show projected face-on and edge-on stellar and gas density maps at $z = 0$ for the galaxies in the star formation efficiency and feedback test suite respectively. The disks show a wide range of spiral galaxy morphologies, and will return to this point in Sect. 6.5. We decompose the resulting stellar disks into a bulge, bar and disk component and fit these simultaneously to the stellar surface density profile. The latter is calculated using all stars out to a height of $|z| = 2.5$ kpc. For the bulge and disk component we assume exponential profiles, i.e.

$$\Sigma(r) = \Sigma_0 \exp(-r/r_\mathrm{d}), \tag{6.9}$$

where we fit for Σ_0 and the scale radius r_d. The bar component is modeled using a simple Gaussian,

$$\Sigma_\mathrm{bar}(r) = \frac{A_0}{\sigma} \exp\left(-\frac{(r - r_0)^2}{2\sigma^2}\right), \tag{6.10}$$

where we fit for the width σ, the central point r_0 and amplitude A_0. We consider this a conservative estimate of the bar mass as a Gaussian contribution falls off towards the center of the disk, leaving more mass to be accounted for by the bulge. An example of the fitting procedure can be seen in Fig. 6.3. The necessity of a separate bar component is here clearly illustrated. In the more bulge dominated cases the bar amplitude is decreased considerably, owing to the weaker disk self-gravity.

The bulge mass, $M_\mathrm{bulge} = 2\pi\Sigma_\mathrm{bulge} r_\mathrm{bulge}^2$, is obtained by integrating Eq. 6.9. A similar relation holds for the disk, where we also include the stellar disk mass past the break radius in the quoted disk stellar mass, $M_\mathrm{disk,s}$, but we only use the data within the break when fitting (see Fig. 6.3). The bar mass is simply the integrated mass from Eq. 6.10, and we consider the bar as a part of the disk component and include it in the quoted

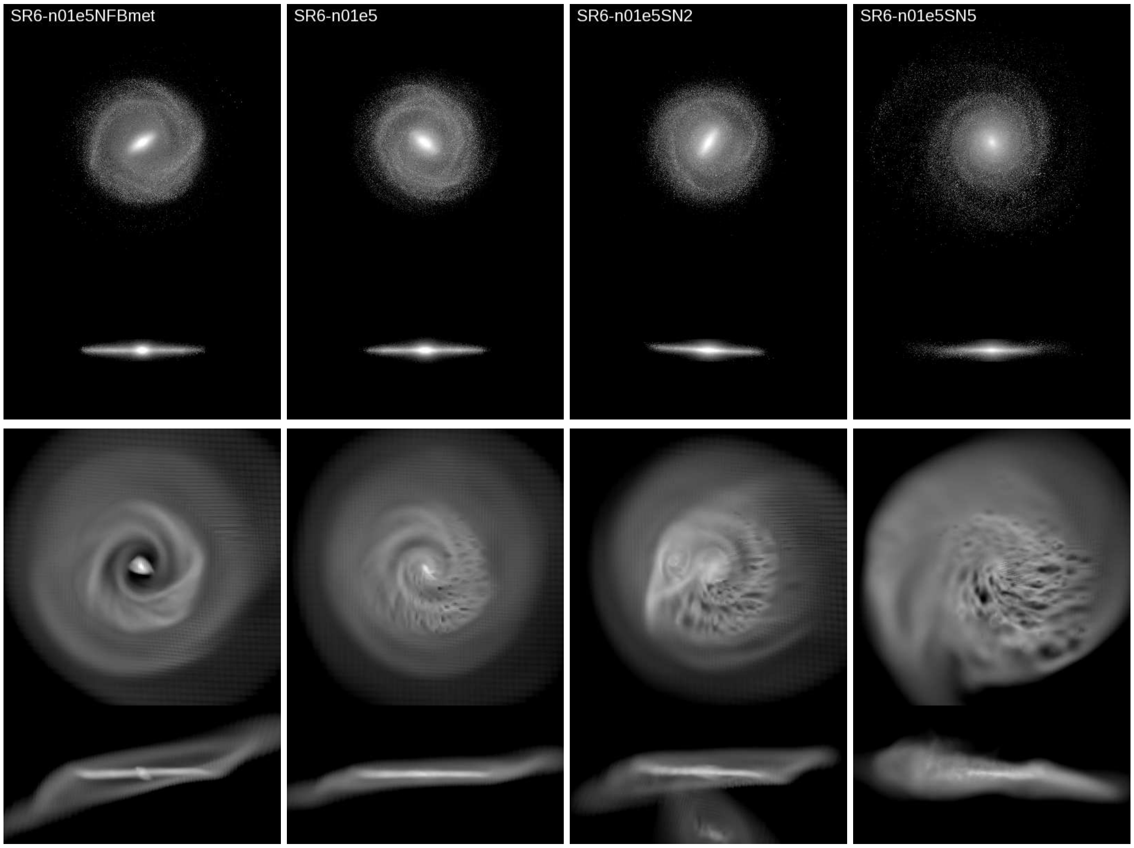

Figure 6.2: Projected face-on and edge-on surface density maps of the stars (top) and gas (bottom) of the $z = 0$ disks where each panel is 60 kpc across. From left to right, the feedback strength is $E_{\rm SNII} = 0, 1, 2$ and $5 \times 10^{51}\,{\rm erg}$. As the injected amount of energy is increased, the gas component becomes more distorted.

$M_{\rm disk,s}$. Doing this or not modifies the disk mass only slightly, especially in bulge dominated galaxies. The gas is treated as a single component, and we simply consider the mass within $r = 15\,{\rm kpc}$ and $|z| = 2.5$ kpc as the gaseous disk mass, $M_{\rm disk,g}$. We consider only the stars when calculating the bulge-to-disk (B/D) and bulge-to-total (B/T) ratios. All measured and derived quantities are summarized in Table 6.2. We note that this method of defining galactic components in simulations, as well as others e.g. via angular momentum (Okamoto et al., 2005; Scannapieco et al., 2009), carry uncertainties.

Our simulated disks span a large range of characteristics: stellar disk masses are in the range $M_{\rm disk,s} = 5 - 9 \times 10^{10}\ {\rm M}_{\odot}$, bulge masses of $M_{\rm bulge,s} = 2 - 7 \times 10^{10}\ {\rm M}_{\odot}$, B/D $\sim 0.23 - 1.2$ and gas fractions $f_{\rm g} = 0.05 - 0.28$. The scale radii of the disks, $r_{\rm d}$, vary from typically $4 - 5$ kpc to > 10 kpc in the bulge dominated systems. As we demonstrate below, extended disk galaxies of Sb, or even Sbc type, form only when star formation is numerically resolved in the whole disk ($n_0 = 0.1\,{\rm cm}^{-3}$), and a low efficiency $\epsilon_{\rm ff} \sim 1$ per cent (or *very* strong feedback) is adopted. At larger efficiencies we observe how the B/D ratios increase, the disks are less extended, the rotational velocities peak at very large values and the spiral patterns become more tightly wound and less pronounced. This indicates a shift towards early type disks like Sa or even S0.

6.5. Effect of Star Formation Parameters

In this section we study the influence of star formation parameters, i.e. in essence the small scale physics, on disk properties at $z = 0$. The resulting stellar surface densities ($\Sigma_{\rm s}$), gas surface densities ($\Sigma_{\rm gas}$) and rotational

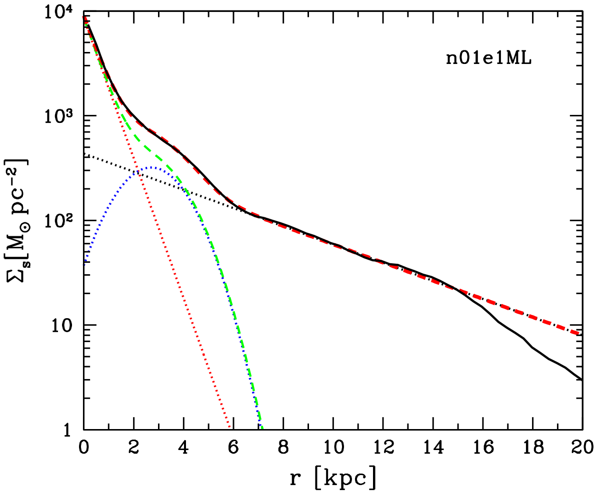

Figure 6.3: Example of a multi-component fit (red dashed line) to the stellar surface density (black solid line) of the $z = 0$ disk in n01e1ML. The fit is composed of a bulge (red dotted line), a bar (blue dotted line) and a disk (black dotted line) component. The green dashed line shows the bulge and bar surface density contribution.

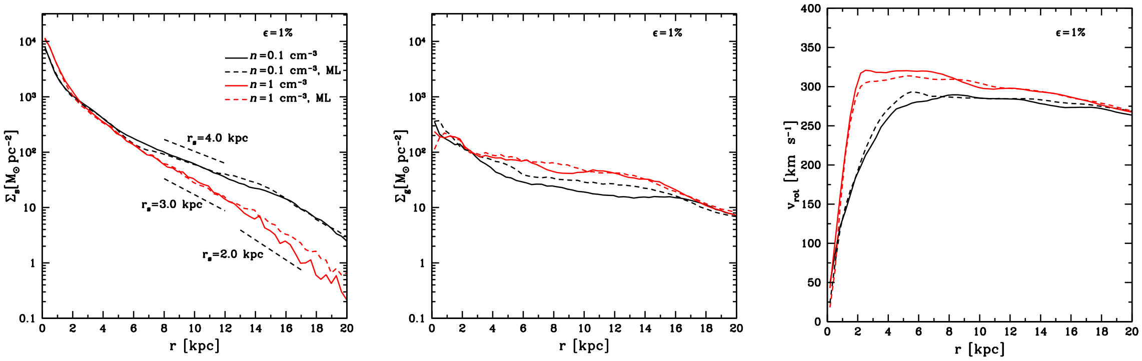

Figure 6.4: The effect of the Schmidt-law star formation density threshold. The panels show the stellar surface density (left), gas surface density (middle) and rotational velocity measured from the gas (right). We consider all material within a height of $|z| < 2.5$ kpc for all components. The star formation efficiency is $\epsilon_{\rm ff} = 1$ per cent in all simulations. The different colours are described in the first panel, and a dashed line indicates that we use the extended feedback model (see text).

velocities measured from the gas ($v_{\rm rot}$) are presented in Fig. 6.4 and Fig. 6.5 for the first 11 simulations in the SR6 suite at $z = 0$ (see Table. 6.1).

The Star Formation Density Threshold, n_0 — We start by focusing on the data presented in Fig. 6.4. By keeping $\epsilon_{\rm ff}$ fixed to 1 per cent, while varying n_0, we observe a strong change in the ability to form stars at large radii. The galaxies adopting a large threshold have a more concentrated distribution of stars, smaller stellar disk scale-lengths as well as larger $\Sigma_{\rm gas}$ at all radii. The scale lengths are $r_{\rm d} > 4,$ kpc for $n_0 = 0.1\,{\rm cm}^{-3}$, but only $r_{\rm d} \sim 2.5\,{\rm kpc}$ for $n_0 = 1\,{\rm cm}^{-3}$. The latter values are on the low side when compared to observations of late type spirals at this mass range (Courteau et al., 1996; Courteau, 1997; Gnedin et al., 2007). The systematically lower $\Sigma_{\rm s}$ signals an under-resolved or "missing" star formation throughout the disk: the average physical gas density does not efficiently cross the targeted n_0, even at intermediate radii. This is also reflected in the gas fractions of ~ 25 per cent, which is much larger than observed average values for

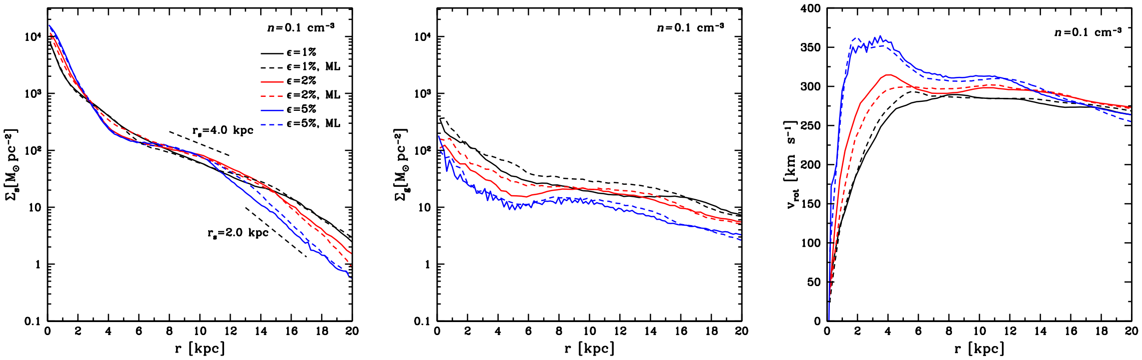

Figure 6.5: The effect of the Schmidt-law star formation efficiency. The panels show the stellar surface density (left), gas surface density (middle) and rotational velocity measured from the gas (right). We consider all material within a height of $|z| < 2.5$ kpc for all components. The Schmidt-law density threshold is $n = 0.1\,\mathrm{cm}^{-3}$ in all simulations. The different colours are described in the first row, and a dashed line indicates that we use the extended feedback model (see text).

galaxies of this size (Garnett, 2002; Zhang et al., 2009). The rotational velocities are large at smaller radii for $n_0 = 1\,\mathrm{cm}^{-3}$, regardless of choice of ϵ_{ff}. Naively one would expect this numerically induced star formation deficiency to alleviate the angular momentum loss at high redshift, hence forming a less concentrated galaxy. However, due to disk secular evolution this is not the case: disk instabilities drive gaseous flows to the center of the disk where, as the gas crosses the correct threshold, star formation can proceed. We conclude, that given our numerical resolution and simulated system, $n_0 = 0.1\,\mathrm{cm}^{-3}$ yields more realistic (average) disk galaxies when compared to observations. We discuss this numerical effect and its relationship to the adopted mesh resolution further in Appendix 6.10. Note that this is not a fundamental result of galaxy formation, but serves only as tuning given our numerical resolution and is a basis for the subsequent tests. A discussed in Sect. 6.2, n_0 should be increased as the resolution is increased.

The Star Formation Efficiency, ϵ_{ff} — We now turn to the data presented in Fig. 6.5, where we keep the threshold fixed at $n_0 = 0.1\,\mathrm{cm}^{-3}$ and adopt $\epsilon_{\mathrm{ff}} = 1, 2$ or 5 per cent. As ϵ_{ff} is increased, Σ_{star} increases at small radii i.e. the bulge mass increases, signaling a lower disk angular momentum. The bulge to disk ratio increases from $\mathrm{B/D} = 0.25$ to 1.25 as ϵ_{ff} increases from 1 to 5 per cent. Σ_{gas} roughly follows a $1/r$-profile and the magnitude is lowered at all radii by approximately the relative change in efficiency. The stellar disk is less extended and the exponential scale-length increases for larger efficiencies (see Table. 6.2). For $\epsilon_{\mathrm{ff}} = 1$ per cent, the disk scale length is measured to be $r_{\mathrm{d}} \sim 4-5\,\mathrm{kpc}$, in good agreement with observed average values from the SDSS (Gnedin et al., 2007), while for $\epsilon_{\mathrm{ff}} = 5$ per cent, $r_{\mathrm{d}} \approx 15\,\mathrm{kpc}$ is a $> 2\sigma$ outlier. Large uncertainties exist for $r_{\mathrm{d}} > 10$ kpc as the stellar disks are small and feature almost flat stellar surface density profiles.

At large radii in all simulations, r_{d} shifts to $\sim 2\,\mathrm{kpc}$. Disk breaking is a well observed phenomenon (Pohlen & Trujillo, 2006) correlated with a dip in the star formation rate, and has been studied numerically by Roškar et al. (2008). As larger star formation efficiencies lead to less extended disks, the disk breaks occurs at smaller radii: $r_{\mathrm{break}} \approx 16, 14, 10$ kpc for $\epsilon_{\mathrm{ff}} = 1, 2$ and 5 per cent respectively. The breaks can also be seen from the average stellar ages $\langle t_* \rangle$, shown in Fig. 6.6. The central parts of the disks generally consist of older stars formed at $z > 1$, and $\langle t_* \rangle$ decreases with larger radii, reaching $\langle t_* \rangle \approx 6\,\mathrm{Gyr}$. Past the disk break, older stars appear which in part can be attributed to stellar migration as well as pollution by old halo stars with $\langle t_* \rangle \approx 11-12\,\mathrm{Gyr}$. We note that $\langle t_* \rangle$ flattens or even declines towards the center of the disks as ϵ_{rmff} is lowered. This is due to secular evolution: as spiral structure is more pronounced (as in n01e1), gas is transported towards the center more efficiently and late time star formation occurs, see Fig. 6.1.

The efficiency has a strong impact on the rotational velocity. The rotation curve in the n01e5 simulation features a strong peak in the inner parts of the disk. As ϵ_{ff} is lowered, B/D decreases and the velocity profile flattens. Only when $\epsilon_{\mathrm{ff}} < 2$ per cent can a flat rotational velocity profile be produced! In n01e1, the rotational

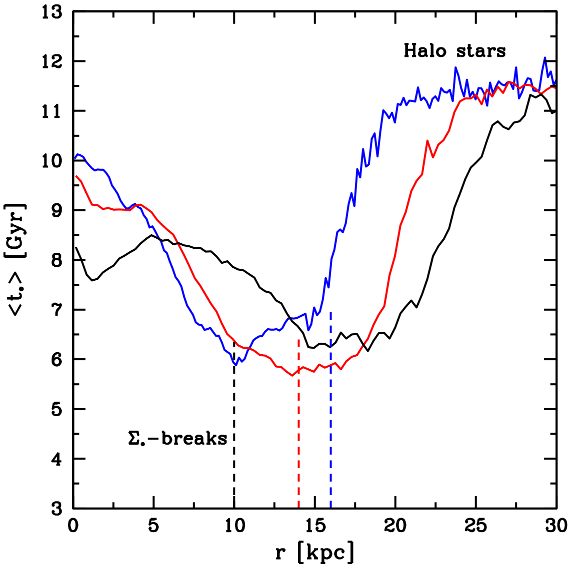

Figure 6.6: Average stellar ages as a function of radius. The colours represent different star formation efficiencies, $\epsilon_{\rm ff}$= 1 (black), 2 (red) and 5 per cent (blue). Dashed lines mark the measured stellar surface density breaks.

velocity reaches $v_{\rm rot} \approx 275\,{\rm km\,s^{-1}}$ and stays roughly flat. For n01e2 and n01e5, $v_{\rm rot}$ peaks at ≈ 310 and $360\,{\rm km\,s^{-1}}$, but converges at $275\,{\rm km\,s^{-1}}$ close to $r = 20$ kpc.

In the left-hand panel of Fig. 6.7, we plot the *circular* velocities, i.e. $v_{\rm c}(r) = \sqrt{GM(<r)/r}$, for the dark matter, disk + bulge baryons and total mass of the galaxies. We note that the $v_{\rm c}$-profiles are well traced by the cold gas rotation curve (the stellar rotational velocities are lower than what can be expected from $v_{\rm c}$ due to a larger velocity dispersion). While all simulations converge at large radii, and have an equal dark matter and baryon contribution within $r \sim 17$ kpc, the mass distribution (and angular momentum distribution) differs dramatically leading to large difference in circular velocities. As we will demonstrate in the next section, a majority of the mass within the bulge component originates from the intense star formation epoch at $z \sim 2-3$ where the value of $\epsilon_{\rm ff}$ matters the most. We also a significantly enhanced dark matter contraction at large $\epsilon_{\rm ff}$. This effect in our simulation suite, and its relevance for direct dark matter detection, has recently been analyzed by Pato et al. (2010).

Star Formation Histories — The left-hand panel of Fig. 6.8 shows the star formation histories for all stars belonging to the disks in n01e1, n01e2 and n01e5 at $z = 0$. The average SFR at the current epoch is $\sim 3-4\ {\rm M_\odot\,yr^{-1}}$ in all simulation, regardless of numerical setting. Moreover, the star formation history during the quiescent phase of disk evolution, i.e. after $z \sim 1$, is relatively flat and roughly the same in all simulations. Significant differences occur at intense epochs of star formation, especially at $z = 3$ where the proto-disk is assembled via cold streams, satellite mergers and gas accretion from the hot halo, as seen in the simulation snapshot in Fig. 6.9[29]. Here the peak of SFR changes dramatically from $\sim 43\ {\rm M_\odot\,yr^{-1}}$ in n01e5 to $\sim 23\ {\rm M_\odot\,yr^{-1}}$ in n01e1.

In n01e5, stars form efficiently everywhere, even in satellites. The gas is quickly consumed locally during the high redshift assembly, and merging systems loose angular momentum to the dark halo, ending up in the central part of the galaxy. Accretion via cold streams from the and hot halo gas will still supply the

---

[29] The image is an RGB composite image where *red* is temperature, *green* is metals and *blue* is density. Each quantity is a mass weighted average along the line of sight. For each image pixel, we calculate the RGB triplet as

$$(R,G,B) = 255\left(\frac{\log(T/T_{\rm m})}{\Delta T}, \frac{\log(Z/Z_{\rm m})}{\Delta Z}, \frac{\log(\rho/\rho_{\rm m})}{\Delta\rho}\right),$$

where $\log\{T_{\rm m}\,[{\rm K}], Z_{\rm m}\,[Z_\odot], \rho_{\rm m}\,[{\rm cm^{-3}}]\} = \{4.1, -3, -4.8\}$ and $\{\Delta T, \Delta Z, \Delta\rho\} = \{2, 2, 6\}$.

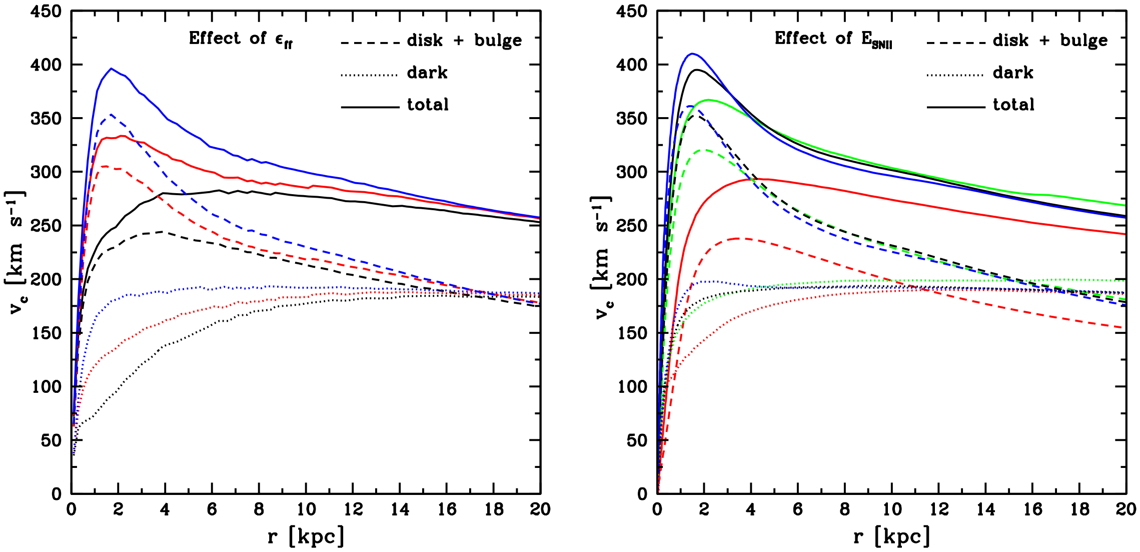

Figure 6.7: Circular velocities, $v_c = \sqrt{GM(<r)/r}$, of the stars (solid), dark matter (dotted) and total mass (dashed). (Left) The effect of the star formation efficiency: $\epsilon_{\rm ff}$= 1 (black), 2 (red) and 5 per cent (blue). (Right) The effect of supernovae feedback strengths (adopting $\epsilon_{\rm ff}$= 5 per cent): $E_{\rm SNII} = 0$ (blue), 10^{51} erg (black), 2×10^{51} erg (green) and 5×10^{51} erg (red).

galaxy with unprocessed gas (Dekel et al., 2009; Agertz et al., 2009), but now in a more bulge dominated environment. In n01e1, star formation is less efficient and a significant portion of the mass in the merging clumps is in gaseous form. This material is lost to the hot gaseous halo via ram-pressure and tidal stripping, or expelled during SNe events, which later cools down to join the disk. Hence, the material that is not consumed by star formation at $z = 3$ is processed at a later epoch, closer to $z = 1 - 2$ (see Fig. 6.8), but now in a more disk like, higher angular momentum configuration. In fact, the trend at $z = 3$ is reversed at this later epoch, and the largest SFR is found for $\epsilon_{\rm ff} = 1$ per cent. These two modes of star formation are related to the classical angular momentum problem (Navarro & White, 1994), and leads to fundamentally different modes of disk assembly and morphology.

A confirmation of the above discussion is shown in Fig. 6.10, where contours of formed stellar masses are outlined in the star formation time-disk radius plane. Note that this mass refers to all the stars at $z = 0$ contained in the disk, and is hence the sum of the stars formed in merging satellites as well as in situ. While the formed stellar mass in n01e1 is smoothly distributed in a roughly exponential profile across the disk at all times, without a clear sign of extreme star formation bursts, the n01e5 simulation shows a strong central concentration of stars formed at $t = 11.5\,{\rm Gyr}$ ($z \sim 3$). This analysis confirms the notion of efficient star forming satellites loosing angular momentum and being dragged into the central parts of the galaxy.

Hubble Types — Fig. 6.1 shows mass weighted projections of the stellar and gas surface densities for n01e1ML, n01e1, n01e2ML, n01e2, n01e5ML and n01e5. The gaseous disks are thin and extended in all simulations, and are surrounded by a warped layer of cold/warm gas, probably associated with misaligned accretion events (Shen & Sellwood, 2006). A hot gaseous halo surrounds the disks, and a temperature projection (not shown) reveals an extended disk-halo interface of warm/hot gas. We will explore this in future work.

As discussed in Sect. 6.5, we find a very strong trend in disk and bulge mass with increasing star formation efficiency. The n01e1 simulation feature a $8.6 \times 10^{10}\,{\rm M}_\odot$ stellar disk with a $2 \times 10^{10}\,{\rm M}_\odot$ bulge, hence B/D$\sim 1/4$. In n01e5 the disk is 35 per cent less massive and the bulge 3.5 times more massive with B/D~ 1.25. Roughly the same scaling holds when including additional SNIa feedback and stellar mass loss.

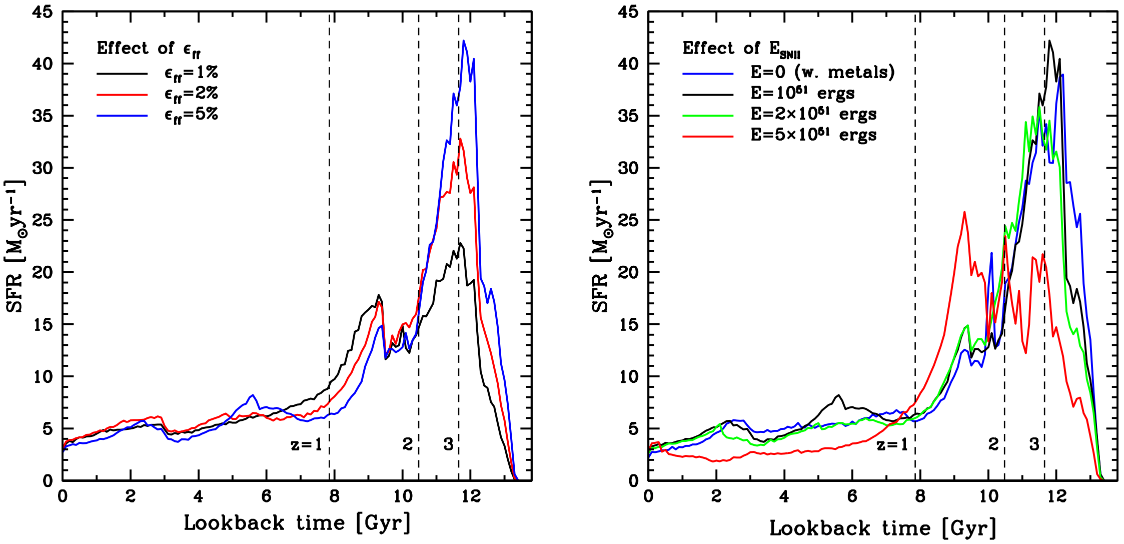

Figure 6.8: (Left) Star formation histories using different values of $\epsilon_{\rm ff}$. At $z < 1$, all simulations regulate to similar SFRs, ending up at $\sim 3 - 4\,M_{\odot}{\rm yr}^{-1}$ at $z = 0$, regardless of star formation parameters. Using a high efficiency leads to central galaxies and dwarfs burning their fuel quickly at high redshift during galaxy assembly, resulting in excess angular momentum loss and a prominent central spheroid. A lower efficiency avoids this issue, leaving more gas left for star formation at lower redshifts in a more disk like configuration. (Right) Star formation histories for a set of simulations of increasing supernovae feedback strength ($E_{\rm SNII}$). We note that the large SFR peak at $z = 3$ is only lowered when a very large amount of energy is injected into the ISM.

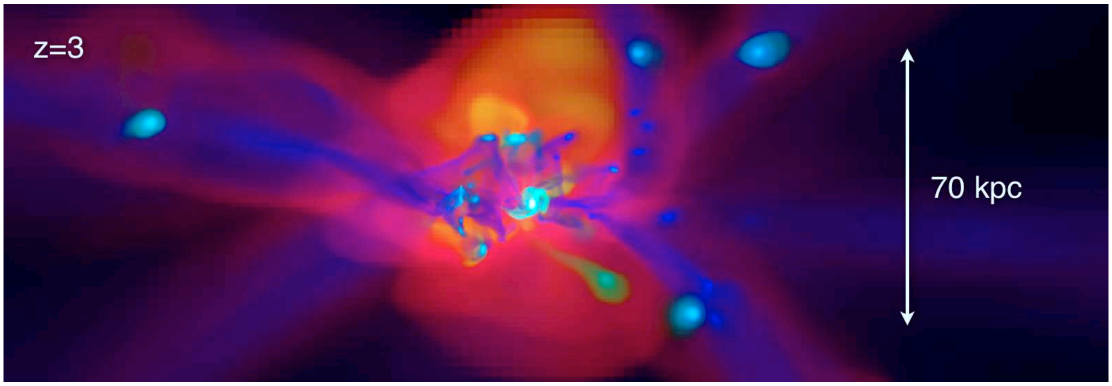

Figure 6.9: A large scale view of the assembling spiral galaxy from the SR5 simulation at $z \sim 3$; the most intense epoch of star formation for this system. The *RGB*-image^2 shows the gas component using temperature (*red*), metals (*green*) and density (*blue*). We can clearly distinguish accretion via streams of cold pristine gas (in blue) penetrating the shock heated gas (in red), reaching the heart of the halo. Dwarf galaxies outside of the large gaseous halo are surrounded by puffy enriched gas originating from stellar outflows. Gas is efficiently lost via tidal and ram-pressure stripping as the dwarfs interact with the main galaxy and its hot gaseous halo. The distance measure is in physical units.

All disks show spiral pattern in the gas component with a larger amplitude in the more gas rich disks, having lower $\epsilon_{\rm ff}$. We also observed spiral structure in the stellar component also, which is the most pronounced in n01e1 and n01e1ML and n01e2ML simulations. As $\epsilon_{\rm ff}$ is increased, B/D increases and the spirals arms become more tightly wound as marginal gravitational instabilities can no longer excite pronounced open spiral arm structure. All disks feature a stellar bar, and viewed edge-on, we observe how the inner stellar

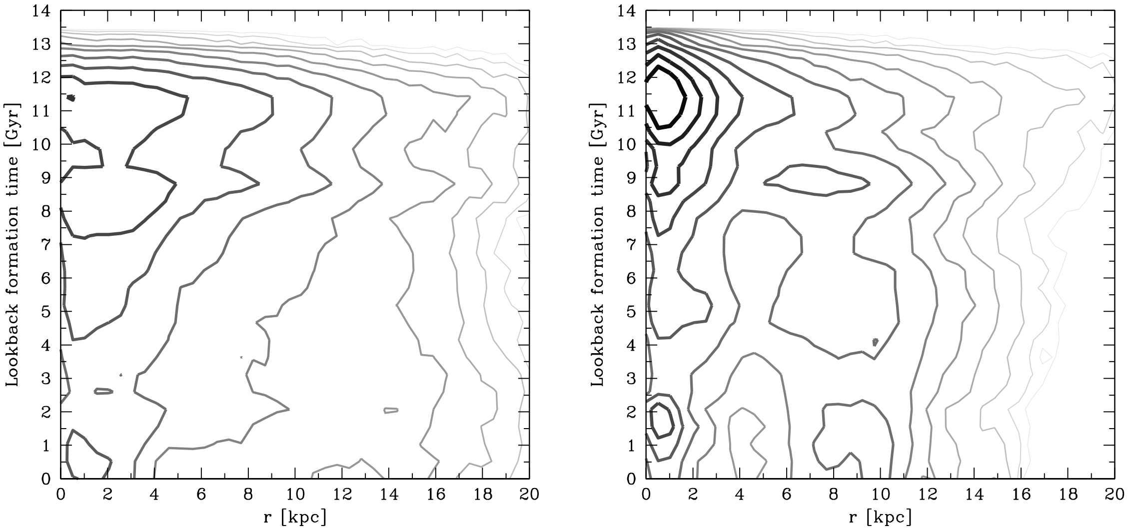

Figure 6.10: Formation time of stars and their radial distribution for the disks in n01e1 (left) and n01e5 (right) at $z = 0$. The contours trace regions of binned mass using bins of size $\Delta t = 0.5\,\mathrm{Gyr}$ and $\Delta x = 0.5\,\mathrm{kpc}$. The contour lines trace, from thin lines with light shades, to thick lines with dark shades, the formed stellar masses from $\log(M_*) = 6.5$ to $\log(M_*) = 9.5$ in steps of 0.25 dex. While the formed stellar mass in n01e1 is smoothly distributed across the disk at all times, the n01e5 simulation shows a strong central concentration of stars formed at $t = 11.5\,\mathrm{Gyr}$ ($z \sim 3$).

distribution flattens as ϵ_{ff} is decreased. The flattened central parts of n01e1 and n01e1ML, and the fact that the bulge is well-fitted using an exponential, is indicative of a bulge formed via secular processes (Kormendy & Kennicutt, 2004) e.g. bar buckling (Debattista et al., 2006). The gaseous bar strengthens at lower ϵ_{ff}, and in n01e1 and n01e1ML gas is transported towards the disk center, triggering star formation. In these simulations, close to 50 per cent of the stars associated with the bulge formed in situ of the disk at $z \lesssim 1$, and only ~ 25 per cent formed at the intense star formation peak at $z \sim 3$. This indicates that a significant portion of the flattened bulge has formed via secular evolution, leading to a pseudo-bulge. This is in stark contrast to the bulge formation epoch seen in the central parts of the n01e5 simulations (right panel of Fig. 6.10), where essentially all bulge stars form at $z \sim 3$.

Weinzirl et al. (2009) (see also Laurikainen et al., 2010), recently analyzed 182 H-band images from the OSUBSGS survey (Eskridge et al., 2002) to obtain B/D and B/T values across the Hubble sequence. Comparing their sample averages (see e.g. their figure 14) to our set of simulations (B/D in Table 6.2) suggests that the final disk in n01e5 is of S0/a type, n01e2 of Sa/Sab type and both n01e1 and n01e1ML of Sb/Sbc type. We consider this agreement only as indicative as each Hubble type spans a wide range of B/D and B/T values. Graham & Worley (2008) presented B/D and B/T flux ratios using a sample of over 400 galaxies observed in the K-band. Their B/D estimates for different Hubble types confirm the classification of our simulated disks. There is no doubt that we are measuring a transformation along the Hubble sequence.

6.6. Effect of Supernova Feedback

As we have demonstrated in the previous sections, a high SFE overproduces the central stellar mass of the galaxy. The inclusion of additional SNIa feedback and stellar mass loss did not drastically change the galaxy properties, even though differences can be seen in Fig. 6.1 (the disks in n01e1 and n01e2 feature much stronger spiral structure) and Fig. 6.4, and more late time star formation is made possible (see Sect. 6.7). The effect of stellar mass loss was studied by Martig & Bournaud (2010) who found a stronger effect on the bulge mass in a similar setting, perhaps due to implementation differences.

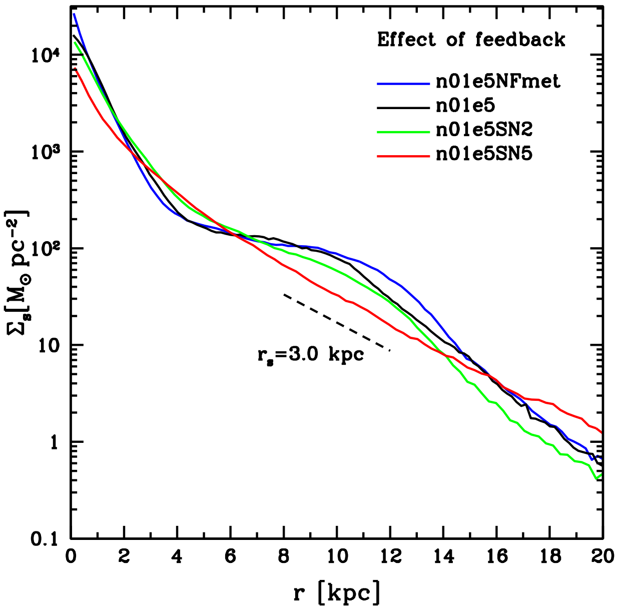

Figure 6.11: Stellar surface densities of the $z = 0$ disks in the supernova feedback test suite. As the SNII feedback energy input is increased, the disk becomes more extended and the bulge component less massive.

In Fig. 6.7 and Fig. 6.11 we present v_c and Σ_s for the n01e5 simulation, but with different amounts of injected SNII energies, see Table 6.2. Note that we still enrich the ISM with metals in the simulation with $E_{\rm SNII} = 0$. Without this, the effect of metal cooling will not be present in all simulations. We find that the bulge mass is lowered as we increase $E_{\rm SNII}$, but only for a very large injected value of 5×10^{51} erg can the disk rotation curve peak at a reasonable $v_c < 300\ \rm km\,s^{-1}$, resembling that of the n01e1 simulation. As for the standard feedback runs in the previous section, the dark matter halo is more contracted as star formation is less regulated. The difference in v_c at $r = 20$ kpc between the n01e5SN5 and the other simulations, corresponding to a few $10^{10}\ \rm M_\odot$, is due to the expelled gas during galaxy assembly which can be accounted for in the more massive gas halo. This effect can also be seen in Σ_s as the central values are decreased, the disks scale radius decreases and the break radius is shifted to larger radii. As seen in Table 6.2, a massive disk still forms. The effect on the star formation histories are shown in the right hand panel of Fig. 6.8; we find no significant difference among the simulations, apart for the very energetic n01e5SN5 simulation. The SFH now resembles that of n01e1 where the $z = 3$ amplitude is lowered to $\approx 20\,\rm M_\odot yr^{-1}$ and more gas is left to form stars in a disk like environment at $z \sim 1 - 2$.

The projected gas density and stellar maps were shown in Fig. 6.2. While the standard feedback simulations shown in Fig. 6.1 showed a clear Hubble sequence of open to tightly wound spiral structure as B/D was lowered, this is not the case for the feedback test suite. In n01e5SN5, the gaseous disk is heavily distorted, warped and puffed up by the large SNII energy injections. Star formation is here very different compared to n01e1 as stars form in filaments and shells from SNe explosions rather than in gas-rich spiral arms.

The effect of metal cooling is not always accounted for in cosmological simulations. Piontek & Steinmetz (2009a) included this effect and reported on difficulties in suppressing the initial high-z peak, even with sophisticated feedback models. In our simulations metal cooling is roughly counter-acted by the standard SNII feedback. If metal enrichment is turned off together with the feedback, we do not find a significant modification to our disks. For example, the n01e1NFB simulation shows a surprisingly successful set of characteristics when compared to n01e1 (see Table 6.2). As a zero metallicity gas cools inefficiently below 10^4 K, as well as in the range 10^5 K to 10^7 K, the n01e1NFB disk essentially behaves as higher metallicity counterpart but with SNII heating balancing cooling. This is the philosophy behind sub-grid multiphase models (e.g. Springel & Hernquist, 2003) in which feedback is implicitly treated as a stiff gas equation of state. Note that our polytropic EOS is slightly stiffer than what is usually adopted ($\gamma = 2$ instead of $\gamma = 5/3$).

6.7. Relationship to Observations

Angular Momentum of the Baryons — For each galaxy we calculate the cumulative specific angular momentum vector, defined as

$$\boldsymbol{j}_{\mathrm{bar}}(\leq r) = \frac{1}{M(\leq r)} \sum_{i=1}^{N(\leq r)} m_i \boldsymbol{x}_i \times \boldsymbol{v}_i, \tag{6.11}$$

including all bulge and disk baryons. Here $\boldsymbol{x_i}$ and $\boldsymbol{v_i}$ are positions and velocities of the gas cells and star particles of the N elements within a radius r encapsulating the mass $M(\leq r)$. The resulting $j_{\mathrm{bar}} = |\boldsymbol{j}_{\mathrm{bar}}|$ are presented in Table. 6.2.

Using the sample of Courteau (1997) and Mathewson et al. (1992), Navarro & Steinmetz (2000) calculated the specific angular momenta vs. rotational velocity for late type spiral galaxies and compared them to numerical simulations. The disks were assumed to follow an exponential profile for which the peak rotational velocity, $v_{\mathrm{rot},2.2}$, occurs at $r = 2.2 r_{\mathrm{d}}$, and it follows that $j_{\mathrm{bar}} = 2 r_{\mathrm{d}} v_{\mathrm{vrot},2.2}$. This assumption can be misleading when comparing to simulated galaxies as the true v_{rot}-peak can be significantly underestimated in the case of bulge-dominated disk galaxies. The sample of Courteau (1997) concerned Sb-Sc galaxies for which B/D is low and a dominating exponential disk assumption is roughly valid. The difference in measured and estimated angular momentum content makes it difficult to compare simulated and observed galaxies, as discussed in Abadi et al. (2003a) and Piontek & Steinmetz (2009b). A simulated galaxy can be considered as a successful realizations of a late type (Sb-Sc/Sd) galaxy if the estimated and measured angular momenta are in agreement.

Focusing on the $n = 0.1\,\mathrm{cm}^{-3}$ suite, we find that the n01e1 and n01e1ML simulations are in good agreement with the observed galaxies, both when analyzed using Eq. 6.11 and the exponential disk approximation. Typical measured and estimated values are here $j_{\mathrm{bar}} \sim 2000\,\mathrm{km\,s^{-1}kpc}$ and $j_{\mathrm{bar}} \sim 2750\,\mathrm{km\,s^{-1}kpc}$ respectively. As ϵ_{ff} is increased, the calculated angular momentum decreases. The $\epsilon_{\mathrm{ff}} = 2$ per cent simulations are still a part of the observed scatter while higher values create more significant outliers in the observed distribution. When using the exponential disk approximation, all simulated galaxies are in good agreement with the observed data as the velocities are quite comparable at larger radii, and for the fact that the disks, although less extended, have larger r_{d} in the higher efficiency cases (see Fig. 6.5). We conclude that an angular momentum reservoir comparable to Sb/Sc galaxies have been reproduced for the baryons in the case of low SFE (i.e. $\epsilon_{\mathrm{ff}} = 1$ per cent).

The lack of correlation between E_{SNII} and the baryonic angular momentum content might come as a surprise. However, while the B/D ratio decreases for large supernova energy injections, the actual disk mass changes little, and is $\sim 6 \times 10^{10}\,\mathrm{M}_{\odot}$ for all $\epsilon_{\mathrm{ff}} = 5$ per cent simulations. As the net contribution of the bulge to the angular momentum content is roughly zero, similar j_{bar} is to be expected. All $\epsilon_{\mathrm{ff}} = 5$ per cent simulations have measured $j_{\mathrm{bar}} \sim 1300 - 1450\,\mathrm{km\,s^{-1}kpc}$ which is close to the estimated $j_{\mathrm{bar}} \sim 1600\,\mathrm{km\,s^{-1}kpc}$ in n01e5SN5.

In summary, the largest measured baryonic specific angular momentum reservoir can be found in simulations using $\epsilon_{\mathrm{ff}} = 1$ per cent due to a massive disk component, regardless of including feedback or not. At higher efficiencies, j_{bar} decreases, again regardless of feedback.

The Tully-Fisher Relationship — The photometric Tully-Fisher (TF) relation (Tully & Fisher, 1977) links the characteristic rotational velocity of a galaxy with its total absolute magnitude. This correlation holds in all typical photometric bands but with variation in functional form (e.g. Pizagno et al., 2007). Early attempts in forming realistic galaxies (e.g. Abadi et al., 2003a) showed off-sets in the observed relation owing to the formation of very concentrated bulge-dominated galaxies with a low star formation activity at late times. Their velocity-magnitude relation had more in common with S0 galaxies (Mathieu et al., 2002). Recent work seems to have improved on these results by SN feedback regulated star formation (Governato et al., 2007; Piontek & Steinmetz, 2009b). These studies place galaxies closer to the observed relation, but this is in part achieved by circumventing the large measured v_{rot} (caused by the dominant bulge) by using the exponential

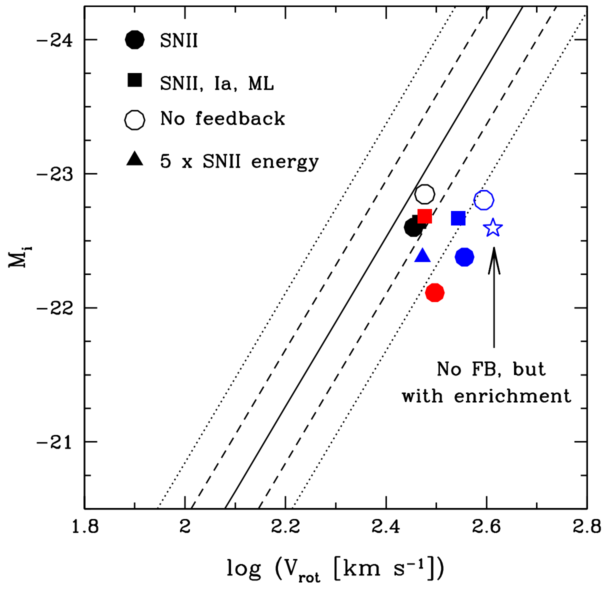

Figure 6.12: The i-band Tully-Fisher relationship from the SDSS (Pizagno et al., 2007). We show the observed average (solid line), 1σ (dashed line) and 2σ (dotted line) relation. The symbols are results from our simulated galaxies, which use $n_0 = 0.1\,\mathrm{cm}^{-3}$ and $\epsilon_{\mathrm{ff}} = 1$ (black symbols), 2 (red symbols) and 5 (blue symbols) per cent.

disk assumption discussed above (but see Governato et al., 2009), i.e. using $v_{\mathrm{rot},2.2}$. Observationally, the measured quantity is often half of the HI velocity width at 20 (W_{20}) or 50 (W_{50}) per cent of the peak intensity.

In Fig. 6.12 we present the measured i-band magnitudes of several of our simulated galaxies as a function of their peak rotational velocities measured from the gas component. These are compared to the observed T-F relation from the SDSS (Pizagno et al., 2007). Pizagno et al. measured the velocity at a radius containing 80 per cent of the i-band flux. This measure (V_{80}) is equivalent to measuring v_{rot} at $\sim 3r_{\mathrm{d}}$ for a pure exponential disk. By using the true peak of v_{rot}, we provide an absolute lower limit to the agreement with observations, and can clearly separate disk and bulge dominated galaxies. We note that the low efficiency models agree well with the average data, regardless of adopted feedback scheme, and even without. At higher efficiencies, the disks are off-set by more than 2σ, mostly due to their peaked rotation curves. In these circumstances, the inclusion of additional recycling via SNIa and stellar mass loss increases the magnitudes by ~ 0.5 dex in n01e2ML and n01e5ML simulations. The n01e5NFB simulation is brighter than the corresponding simulations including feedback due to exclusion of metal enrichment, leading to less efficient cooling and more gas left to form stars at later times. Allowing for enrichment without any energy deposition demonstrates this fact (see figure). From a photometric T-F point of view, the $\epsilon_{\mathrm{ff}} = 5$ per cent disks correspond to S0 systems or early type spirals (Mathieu et al., 2002).

As described in the previous sections, v_{rot} and Σ_{s} in the n01e5SN5 simulation agrees fairly well with the disk values found in n01e1. The strong feedback brings the galaxy closer to the observed values but the absolute magnitude is still lower than in the $\epsilon_{\mathrm{ff}} = 1$ per cent simulations. The SFH in Fig. 6.8 tells us why: after $z = 1$ the SFR is lower in n01e5SN5 compared to n01e1 by almost a factor of 2 (even though the $z = 0$ values agree) due to strong gas expulsion, resulting in a less bright disk by $\sim 1/3$ dex in i-band magnitude.

As for the specific angular momentum analysis, adopting the $v_{\mathrm{rot},2.2}$ measure (or V_{80}), *all* disks would agree statistically with the observed T-F relation, especially when including SNIa feedback and stellar mass loss.

Similar to the photometric T-F is the “baryonic T-F relation" (McGaugh et al., 2000, 2010) which links characteristic rotation velocity with total galaxy baryonic mass. The baryonic TF relation therefor accounts for the fact that less massive galaxies are more gas rich, and their stars only account for a small fraction of the total disk mass. The same conclusion as above holds for the baryonic T-F: the low efficiency simulations agrees well with the observations. As the baryonic masses for the disks are not strongly affected, even in the case of extreme feedback (n01e5SN5) the data points shift only with the increase of the $v_{\rm rot}$ peak. As for the photometric T-F, using $v_{\rm rot,2.2}$ puts all galaxies on the observed relation.

The $\Sigma_{\rm SFR}$-$\Sigma_{\rm gas}$ Relation — The most famous study of the globally averaged relationship between the star formation rate and gas surface density is from Kennicutt (1998) (from now on K98), where a sample of 61 nearby normal spiral galaxies and 36 infrared-selected starburst galaxies were considered. Assuming a Schmidt-law of the form

$$\Sigma_{\rm SFR} = a \left(\frac{\Sigma_{\rm gas}}{1\,{\rm M_\odot\,pc^{-2}}} \right)^N , \tag{6.12}$$

the full sample yielded $a = (2.5 \pm 0.7) \times 10^{-4}$ and $N = 1.40 \pm 0.15$. The $\Sigma_{\rm SFR}$-$\Sigma_{\rm gas}$ relation has been studied by many authors (e.g. Wong & Blitz, 2002; Misiriotis et al., 2006; Kennicutt et al., 2007; Schuster et al., 2007), both locally and globally, using different star formation tracers and galaxy samples. A large range of power-law indices ($N \approx 1-3$) have been found, suggesting that either different SF laws exist in different galaxies or that N is very sensitive to systematic differences in methodology. Bigiel et al. (2008) presented a comprehensive analysis of the $\Sigma_{\rm SFR}$-$\Sigma_{\rm gas}$ relationship using multifrequency data of 7 spiral galaxies and 11 late-type and dwarf galaxies. The analysis pointed to a great variation within the sample and a markedly different functional behaviour in atomic and molecular dominated gas.

The THINGS data of Bigiel et al., relevant for spirals, as well as the K98 law (Eq. 6.12), is reproduced in the left hand panel in Fig. 6.13 together with the azimuthally averaged ($\Delta r = 540\,{\rm pc}$) data from n01e1, n01e2 and n01e5. For the calculation of $\Sigma_{\rm SFR}$ we only consider stars younger than 50 Myr. At a given value of $\Sigma_{\rm gas}$ we find a clear trend of higher $\Sigma_{\rm SFR}$ values for higher $\epsilon_{\rm ff}$. All simulations fall onto the range of observed values, having the same functional behaviour but with an off-set. We note that only the disk in the n01e5 simulations is compatible with the K98 relation. The n01e1 simulation is on the low side but can still statistically be associated with one of the THINGS spiral galaxies. However, at high redshift the argument can be reversed, as can be seen in the right hand panel of Fig. 6.13. The observations of DLAs at $z \sim 3$ by Wolfe & Chen (2006) are typically an order of magnitude lower than the K98 relation, agreeing only with measurement of the low density environment of the disks in our low efficiency simulations. This trend is also predicted by simulations including treatment of H_2 formation (Gnedin & Kravtsov, 2010b,a). In essence, while n01e1 is on the low side at $z = 0$, it is consistent with high redshift observations and the reverse argument is valid for n01e5. A higher efficiency is acceptable at lower redshift, and is predicted due to e.g. higher gas metallicity. As the bulge component is assembled at high redshift, the efficiency of star formation during this epoch is crucial in setting the morphology of the galaxy.

The same analysis is performed for the feedback test suite in Sect. 6.6, and shown in Fig. 6.14. At $z = 0$, all simulations show a similar functional behaviour, but with a weak trend of lower $\Sigma_{\rm SFR}$ as $E_{\rm SNII}$ is increased, while remaining comparable to the K98 law. At $z = 3$, a slightly greater effect is found, but only for very large energy injections ($E_{\rm SNII} \geq 2 \times 10^{51}\,{\rm erg}$). The extreme case of $E_{\rm SNII} = 5 \times 10^{51}\,{\rm erg}$ (n01e5SN5) is comparable to a lowering star formation efficiency to $\epsilon_{\rm ff} = 2$ per cent (n01e2). None of the strong feedback simulations regulate star formation enough to reproduce the low $\Sigma_{\rm SFR}$ values found for $\epsilon_{\rm ff} = 1$ per cent (n01e1). This $z \sim 3$ insensitivity of the K-S relation to feedback was also found by Kravtsov (2003).

6.8. Discussion and Conclusions

In this paper we have presented a set of Adaptive Mesh Refinement simulations studying the assembly of large Milky Way-like disk galaxies. The self-consistent formation of a late type disk galaxy has remained elusive in the field of numerical galaxy formation, mainly due to the strong loss of angular momentum in the galaxy assembly process. A popular solution to this problem is to regulate star formation at high redshift via supernova explosions that drive galactic winds, transporting material out of star forming regions hence lowering the local star formation rate.

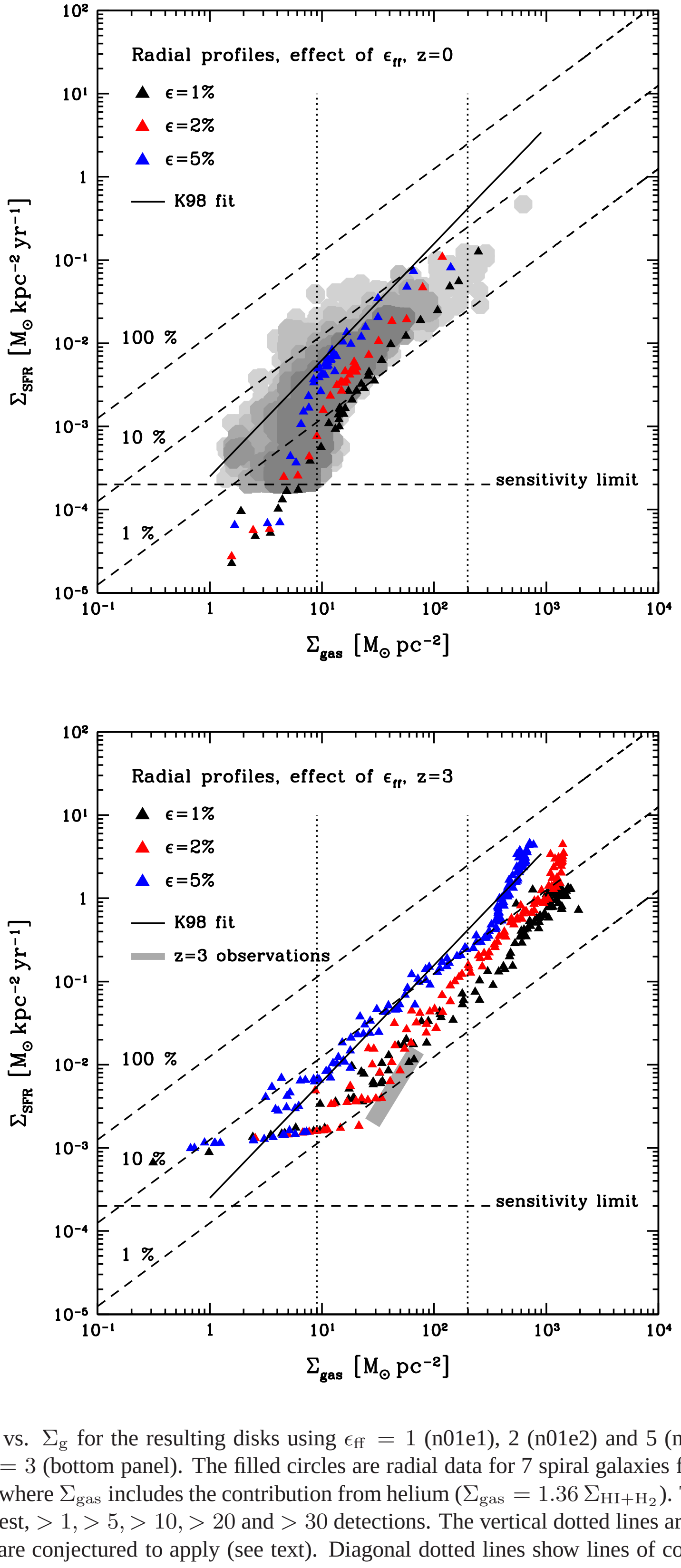

Figure 6.13: $\Sigma_{\rm SFR}$ vs. $\Sigma_{\rm g}$ for the resulting disks using $\epsilon_{\rm ff} = 1$ (n01e1), 2 (n01e2) and 5 (n01e5) per cent at $z = 0$ (top panel) and at $z = 3$ (bottom panel). The filled circles are radial data for 7 spiral galaxies from the THINGS survey (Bigiel et al., 2008), where $\Sigma_{\rm gas}$ includes the contribution from helium ($\Sigma_{\rm gas} = 1.36\,\Sigma_{\rm HI+H_2}$). The data points represent, from lightest to darkest, $> 1, > 5, > 10, > 20$ and > 30 detections. The vertical dotted lines are regions where different star formation laws are conjectured to apply (see text). Diagonal dotted lines show lines of constant SFE=$\Sigma_{\rm SFR}/\Sigma_{\rm gas}$, indicating the level of $\Sigma_{\rm SFR}$ needed to consume 1, 10 and 100 per cent of the gas reservoir in 10^8 years. The solid black line is the average relation from Kennicutt (1998). The $z = 3$ observations approximately populate the region of the Wolfe & Chen (2006) observations. THINGS data courtesy of F. Bigiel.

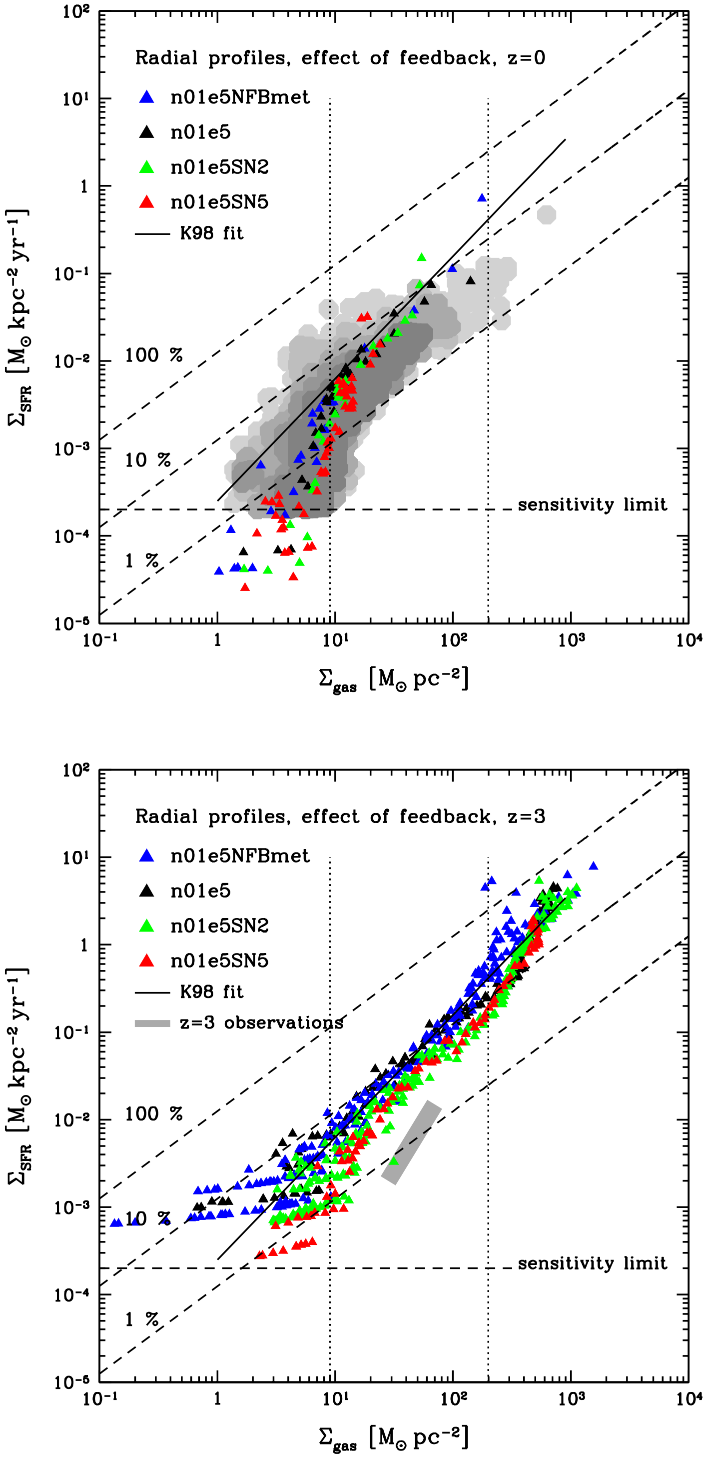

Figure 6.14: $\Sigma_{\rm SFR}$ vs. $\Sigma_{\rm g}$ for the resulting disks using a high star formation efficiency ($\epsilon_{\rm ff} = 5$ per cent), but with different SN feedback strengths: $E_{\rm SNII} = 10^{51}$ erg (n01e5), 2×10^{51} erg (n01e5SN2), 5×10^{51} erg (n01e5SN5) as well as with zero SNII feedback energy but metal enrichment (n01e5NFBmet). The panels show the results at $z = 0$ (top panel) and at $z = 3$ (bottom panel). The lines and symbols are described in the caption of Fig. 6.13. SN feedback has little effect on the $\Sigma_{\rm SFR}$-$\Sigma_{\rm g}$ relation at $z = 0$ but does affect the high redshift relation, although only for very large energy injections ($E_{\rm SNII} \geq 2 \times 10^{51}$ erg).

We have investigated the plausibility of this mechanism in comparison to a small scale (~ 100 pc) physical approach where star formation is made inefficient by modifying the Schmidt-law star formation normalization. In a very crude way, this mimics unresolved physics such as H_2 formation, small scale turbulence and radiative effects. We find that the Schmidt-law efficiency of star formation is far more successful way of regulating star formation towards realistic galaxies than what can be achieved via supernova feedback. Our most successful models reproduce Milky Way galaxies with flat rotation curves, where the small bulge component is formed via secular processes. The main conclusions of this work can be summarized as:

- Disk characteristics such as $\Sigma_*(r)$, $\Sigma_{\rm gas}(r)$, $v_{\rm rot}(r)$ and B/D strongly depends on the choice of star formation efficiency per free fall time, $\epsilon_{\rm ff}$. The parameter will essentially set the mode of global star formation, hence governing the final spiral Hubble type, where low efficiencies of $\epsilon_{\rm ff} \sim 1$ per cent render disks of Sb or Sbc type, while $\epsilon_{\rm ff} = 5$ per cent moves the disks closer to Sa/S0 types. Simulations at low efficiencies agree well with observational constraints on disk characteristics (Courteau, 1997; Gnedin et al., 2007), as well as the angular momentum content of disk galaxies (Navarro & Steinmetz, 2000), the Tully-Fisher relationship (Pizagno et al., 2007) and the $\Sigma_{\rm SFR}$-$\Sigma_{\rm gas}$ relation (Kennicutt, 1998; Bigiel et al., 2008). The origin of the successful Milky Way-like galaxy formation is a well motivated suppression of star formation at $z \sim 3$, the epoch at which the violent assembly process would form a slowly rotating bulge in case of efficient star formation.
- Supernova feedback does not regulate star formation efficiently at low input energies. Only when the injected energy per supernova event is 5 times the canonical value, i.e. $5 \times 10^{51}\,\rm erg$, do we find lower and more realistic B/D ratios in the simulations tuned to the standard Kennicutt (1998) star formation law, leading to a flatter rotational velocity profile, hence resembling the galaxies formed without strong feedback but with a low Schmidt-law efficiency. This comes at the cost of a significantly distorted gas disk at $z = 0$, as well as a less bright stellar disk as gas is expelled into the halo, leaving less fuel for star formation at late times. In essence, we find that changes in $\epsilon_{\rm ff}$ can play a much greater role in shaping a spiral galaxy than gas redistribution via supernovae driven winds.

  It is plausible that at very high resolution, or using a drastically different recipe of supernovae feedback, lower values of $E_{\rm SNII}$ may be successful in regulating the SFE. If so, it will still need to mimic the low efficiency on scales of a few $100\,\rm pc$ which, as argued in this work, can be absorbed by the $\epsilon_{\rm ff}$-term.
- If the star formation efficiency parameter is tuned to match the standard $z = 0$ K-S data (Kennicutt, 1998), i.e. requiring on the order of $\epsilon_{\rm ff} \geq 5\,\rm percent$ (e.g. Stinson et al., 2006), star formation is likely to be overestimated at high redshift ($z = 3$) where the amplitudes of $\Sigma_{\rm SFR}$ are an order of magnitude lower (Wolfe & Chen, 2006; Gnedin & Kravtsov, 2010b). All efficiencies studied in this work ($\epsilon_{\rm ff} = 1 - 5$ per cent) are compatible with modern data of the THINGS survey (Bigiel et al., 2008) but only when $\epsilon_{\rm ff} \sim 1$ per cent can the constraints from $z = 3$ data be met and late-type, disk dominated systems form. As the true SFE varies in space and time, being dependent on small scale physics governing H_2 formation (see e.g. Gnedin et al., 2009), present day simulations based on single valued efficiency parameter have little predictive power.

We argue (see also Gnedin et al., 2009) that the results presented in this paper indicate that other processes in the ISM in addition to, or in conjunction with, supernova feedback are important in explaining the evolution of the galaxy population, as well as regulating observed disk sizes. Some form of outflow process must be responsible for enriching the IGM (Oppenheimer & Davé, 2006), which together with an inefficient star formation might explain the faint end of the stellar mass function (Somerville & Primack, 1999; Kereš et al., 2009). The same argument can be used for the mass-metallicity relationship (Brooks et al., 2007), although Tassis et al. (2008) demonstrated that it could be reproduced without supernova-driven outflows. Galaxies of masses considered in this work are situated at the knee of the stellar mass function, where the observed and simulated functions (even without feedback, see Kereš et al., 2009) are in closest agreement. This circumstance might explain why even our simulations without feedback resulted in realistic disks. At this galaxy mass, supernova driven winds cannot escape the deep potential well, and are impeded by the hot halo. On the other hand, AGN feedback, which recently has been introduced into galaxy formation simulations (Di Matteo et al., 2005), is probably not relevant for the Milky Way since the black hole might not be massive enough for efficient AGN radio-heating. At higher masses, and/or at high redshift, the inclusion of AGN is probably necessary to correctly reproduce the observed abundances and stellar masses. This is the greatest uncertainty of our work, which we leave for a future study.

The way in which galaxies populate dark matter haloes is an important topic, see e.g. Dutton et al. (2010) and references within for a compilation of recent observational data and theoretical work. Recently, Guo et al. (2010) (see also Moster et al. (2010) and Behroozi et al. (2010)) matched dark matter halo mass function from cosmological $N-$body simulations to the stellar mass function of the galaxies from the SDSS (Li & White, 2009). This analysis yields the required galaxy formation efficiency, $\eta = (M_*/M_{\rm halo})(\Omega_{\rm m}/\Omega_{\rm b})$, i.e. what fraction of the universal baryons that must have condensed into stars at a given halo mass. In our "best-case" model (n01e1ML, see Table 6.1), the total stellar and dark matter halo virial mass is $\sim 10^{11}$ $\rm M_{\odot}$ and $\sim 10^{12}$ $\rm M_{\odot}$ respectively. This results in a stellar fraction of 10 per cent, which corresponds to almost 60 per cent of the cosmic baryon fraction. The rest of the baryons reside in the stellar halo, gaseous disk and ionized gas halo. At this halo mass, abundance matching requires that the stellar disk accounts for only ~ 20 per cent of the cosmic baryon fraction, i.e. a factor of three lower. Similar discrepancies exist in all modern work of numerical galaxy formation (Abadi et al., 2003a; Okamoto et al., 2005; Governato et al., 2007; Scannapieco et al., 2009; Piontek & Steinmetz, 2009b), and its origin is not yet know, although AGN is a compelling mechanism at the high mass end, as discussed above. This issue is the topic of a follow-up paper in preparation.

Behroozi et al. (2010) performed a comprehensive analysis of abundance matching, accounting for systematic errors in e.g. the stellar mass estimates, the halo mass function, cosmology etc. Our simulated galaxy formation efficiencies would be in $\sim 2\sigma$ agreement with their result (see their Fig. 11). We note that our own Galaxy and M31 also might be strong outliers in this analysis, considering the inferred η from mass modeling (Klypin et al., 2002; Seigar et al., 2008) as well as via recent MW halo mass estimates (Xue et al., 2008).

Abundance matching is insensitive to the actual Hubble types of the galaxies, forcing all galaxies of a specific mass to be linked to only one halo mass. At the stellar mass scale of the Milky Way ($5-7\times10^{10}$ $\rm M_{\odot}$), only ~ 25 per cent of galaxies are of Sb/Sbc type (Nair & Abraham, 2010). It is plausible that the more active merger histories associated with ellipticals and early type disk galaxies have led to a stronger mass expulsion, via e.g. AGN, in comparison to the more disk dominated counterparts. In this scenario, late type disks are expected to be outliers in the galaxy formation efficiency vs. stellar mass relation, considering the strong bias towards early type systems. A detailed sub-division into Hubble types has not yet been performed when matching galaxies to haloes, although color separations into red and blue systems have been made in studies using weak-lensing (Mandelbaum et al., 2006) and satellite kinematics (More et al., 2010). These studies indicate a different galaxy formation efficiency for galaxies similar in mass to the Milky Way; a late type galaxy is associated with a halo of ~ 0.5 dex lower halo mass compared to an equally massive early type (see e.g. Fig. 11 in More et al. (2010)).

Understanding, from a numerical perspective, the spread of baryon fractions across dark matter haloes of different masses, accretion histories and environments is a complicated problem, and will require a large sample of high-resolution simulations, which we leave for a future investigation.

Acknowledgments

We are very grateful to D. Potter for generating the initial conditions used in this paper. We thank F. Bigiel for providing a copy of his data. We thank Andrey Kravtsov, Joseph Silk, Simon White, Joop Schaye, Francois Hammer and Alister Graham for valuable comments. This work was granted access to the HPC resources of CINES and CCRT under the allocation 2009-SAP2191 made by GENCI (Grand Equipement National de Calcul Intensif). We have also made use of the zBox3 and Schrödinger supercomputers at the University of Zürich.

6.9. Effect of Resolution

The numerical setup in the SR5-n1e1ML simulation is identical to SR6-n1e1ML, apart from having twice as high spatial resolution and hence eight times the mass resolution (see Table 6.1). In Fig. 6.15 we plot the resulting stellar surface densities and circular velocities from the simulations. We find very little difference

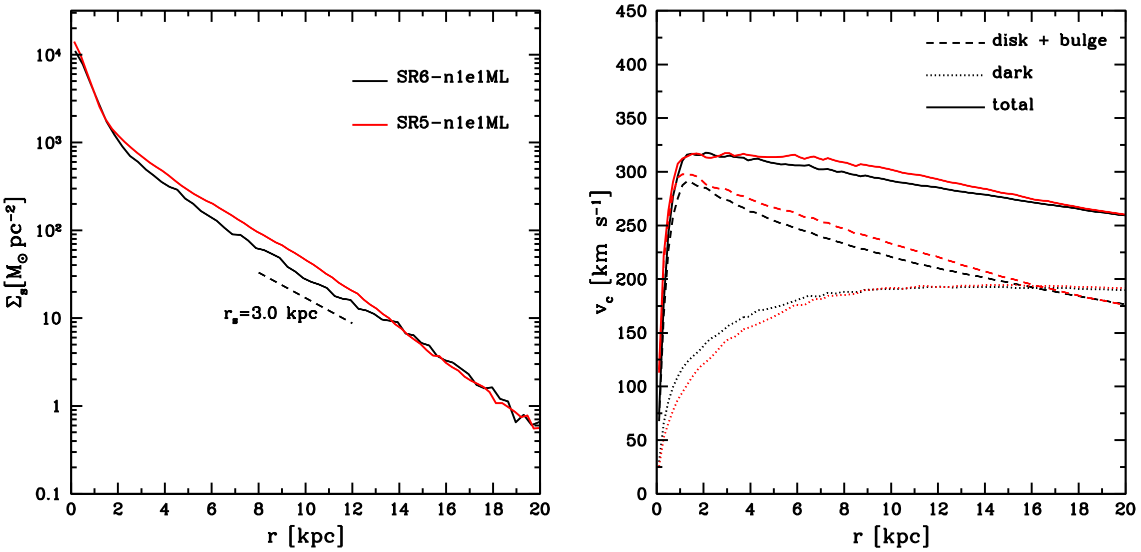

Figure 6.15: Stellar surface densities (left) and circular velocities (right) in SR6-n1e1ML (black line) and SR5-n1e1ML (red line).

apart from more gas being consumed in SR5-n1e1ML resulting in a more massive disk (by ~ 35 per cent) at higher resolution and the bulge mass is lowered by ~ 20 per cent. The dark halo is less contracted but as $M(< r)$ remains roughly the same, the circular velocities changes little.

Why do we not see any striking differences in the higher resolution simulation? At the adopted resolutions we do not uncover more physics with only a factor of two increase in spatial resolution: we are still under-resolving the scale height and are not resolving individual star forming clouds. At much higher resolution this conclusion will change and the numerical parameters will have a different meaning.

6.10. Resolving Star Formation

In this appendix we derive a simple relationship of the required resolution (Δx) and star formation density threshold (n_0) required to capture star formation in an extended disk at the current epoch. Let us assume that a gas disk at $z = 0$ follows an exponential radial density profile with a sech^2 vertical profile, i.e.

$$\rho(r, z) = \rho_0 \,\mathrm{sech}^2(-z/h) \exp(-r/r_d). \tag{6.13}$$

The mass of this profile can be integrated from Eq. 6.13 to give us the characteristic scale disk density

$$\rho_0 = M_\mathrm{d}/4\pi h\, r_d^2. \tag{6.14}$$

Assume that this density distribution is coarsened on a regular mesh with a cell size Δx, and that the disk is aligned to a mesh axis. The physical density in a central strip of cells can then be calculated as

$$\rho(r)_{\Delta x} = \frac{M_\mathrm{d}}{4\pi h\, r_\mathrm{d}^2 \Delta x} \exp(-r/r_d) \left(\int_{-\Delta x/2.0}^{\Delta x/2.0} \mathrm{sech}^2(-z/h) \mathrm{d}z \right) \tag{6.15}$$

$$= \frac{M_\mathrm{d}}{4\pi h\, r_d^2 \Delta x} \exp(-r/r_d) h \left[-\tanh(-z/h) \right]_{-\Delta x/2.0}^{\Delta x/2.0} \tag{6.16}$$

The true star-forming scale height of cold molecular gas is on the order of $\sim 10\,\mathrm{pc}$, and for the neutral atomic ISM it is $\sim 100\,\mathrm{pc}$. Galaxy formation simulations under-resolve this structure, and it is safe to assume that the effective numerical scale height $h \sim \Delta x$, until Δx goes below a few tens of parsecs. Under this assumption the integration limits in the sech^2 term above will always be from $-z/h = -0.5$ to $z/h = 0.5$. Eq. 5 can now be simplified to

$$\rho(r)_{\Delta x} = 0.924 \frac{M_{\mathrm{d}}}{4\pi h\, r_d^2 \Delta x} \exp(-r/r_d) \tag{6.17}$$

which in units of particles per cm^{-3} this is approximately

$$n(r)_{\Delta x} \approx \left(\frac{M_{\mathrm{d}}}{10^9\,\mathrm{M}_\odot} \right) \frac{3 \exp(-r/r_d)}{\Delta x\, r_d^2} \quad \mathrm{cm}^{-3} \tag{6.18}$$

We can rearrange this equation and write down the following useful relationship:

$$r = r_{\mathrm{d}} \ln \left[\left(\frac{M_{\mathrm{d}}}{10^9\,\mathrm{M}_\odot} \right) \frac{3}{\Delta x\, r_{\mathrm{d}}^2 n_{\Delta x}} \right]. \tag{6.19}$$

This equation tells us at which radius, given a cell size, we cross the density $n_{\Delta x}$. In Sect. 6.5 we argued that using $n = 1\,\mathrm{cm}^{-3}$ and $\Delta x = 340$ pc resulted in less realistic galaxies as star formation was "missing" at large radii. Inserting these values into Eq. 6.19 gives us $r \sim 6 - 8\,\mathrm{kpc}$ for any reasonable choice of r_{d}, confirming the arguments of Sect. 6.5. Adopting $n = 0.1\,\mathrm{cm}^{-3}$ increases the star forming radius by a factor of 2.3.

7.
Collaborations and the Future

7.1. Collaborative Projects

This section briefly summarizes the collaborative projects I have been involved in during my doctoral studies. A brief abstract is presented for each project together with its associated publication reference.

Quaquaversal Pre-Initial Conditions for N-Body Simulations — N-body simulations sample their initial conditions on an initial particle distribution, which for cosmological simulations is usually a glass or grid, whilst a Poisson distribution is used for galaxy models, spherical collapse etc. These pre-initial conditions have inherent correlations, noise due to discreteness and preferential alignments, whilst the glass distribution is poorly defined and computationally expensive to construct. We present a novel particle distribution which can be useful as a pre-initial condition for N-body simulations, using a simple construction based on a “quaquaversal” tiling of space. This distribution has little preferred orientation (i.e. is statistically isotropic), has a rapidly vanishing large scale power-spectrum ($P(k) \sim k^4$), and is trivial to create. It should be particularly useful for warm dark matter and cold collapse simulations.

Publication: Hansen et al. (2007)

The Source of Ionization along the Magellanic Stream — Since its discovery in 1996, the source of the bright Hα emission (up to 750 mR[30]) along the Magellanic Stream has remained a mystery. There is no evidence of ionising stars within the HI stream, and the extended hot halo is far too tenuous to drive strong shocks into the clouds. We now present a hydrodynamical model that explains the known properties of the Hα emission and provides new insights on the lifetime of the Stream clouds. The upstream clouds are gradually disrupted due to their interaction with the hot halo gas. The clouds that follow plough into gas ablated from the upstream clouds, leading to shock ionisation at the leading edges of the downstream clouds. Since the following clouds also experience ablation, and weaker Hα ($100-200$ mR) is quite extensive, a disruptive cascade must be operating along much of the Stream. In our model, the clouds are evolving on timescales of $100-200$ Myr, such that the Stream must be replenished by the Magellanic Clouds at a fairly constant rate. The ablated material falls onto the Galaxy as a warm drizzle which suggests that diffuse ionized gas at 10^4K may be an important constituent of galactic accretion. The observed Hα emission provides a new constraint on the rate of disruption of the Stream and, consequently, the infall rate of metal-poor gas onto the Galaxy. When the ionized component of the Stream is fully accounted for, the rate of gas accretion is 0.4 $M_\odot$ yr^{-1}, roughly twice the rate deduced from HI observations alone.

Publication: Bland-Hawthorn et al. (2007)

Thin, Thick and Dark Discs in ΛCDM — In a ΛCDM cosmology, the Milky Way accretes satellites into the stellar disc. We use cosmological simulations to assess the frequency of near disc plane and higher inclination accretion events, and collisionless simulations of satellite mergers to quantify the final state of the accreted material and the effect on the thin disc.

---

[30] 1 Rayleigh (R) = $10^6/4\pi$ photons cm^{-2} s^{-1} sr^{-1}, equivalent to 5.7×10^{-18} erg cm^{-2} s^{-1} arcsec^{-2} at Hα.

On average, a Milky Way-sized galaxy has 3 subhalos with $v_{\max} > 80$ km/s; 7 with $v_{\max} > 60$ km/s; and 15 with $v_{\max} > 40$ km/s merge at redshift $z \gtrsim 1$. Assuming isotropic accretion, a third of these merge at an impact angle $\theta < 20^{\circ}$ and are dragged into the disc plane by dynamical friction. Their accreted stars *and dark matter* settle into a thick disc. The stellar thick disc qualitatively reproduces the observed thick disc at the solar neighbourhood, but is less massive by a factor $\sim 2 - 10$. The dark matter disc contributes $\rho_{\mathrm{DDISC}} = 0.25 - 1\,\rho_{\mathrm{HALO}}$ at the solar position. Although not likely to be dynamically interesting, the dark disc has important implications for the direct detection of dark matter because of its low velocity with respect to the Earth.

Higher inclination encounters $\theta > 20^{\circ}$ are twice as likely as low inclination ones. These lead to structures that closely resemble the inner/outer stellar halos recently discovered by Carollo et al. (2007). They also do more damage to the Milky Way stellar disc creating a more pronounced flare, and warp; both long-lived and consistent with current observations. The most massive mergers ($v_{\max} \gtrsim 80$ km/s) heat the thin disc enough to produce a thick disc. These heated thin disc stars are essential for obtaining a thick disc as massive as that seen in the Milky Way; they likely comprise some $\sim 50 - 90\%$ of the thick disc stars. The Milky Way thin disc must reform from fresh gas after $z = 1$.

Only one in four of our sample Milky Way halos experiences mergers massive and late enough to fully destroy the thin disc. We conclude that thick, thin and dark discs occur naturally within a ΛCDM cosmology.

Publication: Read et al. (2008)

Discreteness Effects in Lambda Cold Dark Matter Simulations: A Wavelet-Statistical View — The effects of particle discreteness in N-body simulations of Lambda Cold Dark Matter (ΛCDM) are still an intensively debated issue. In this paper we explore such effects, taking into account the scatter caused by the randomness of the initial conditions, and focusing on the statistical properties of the cosmological density field. For this purpose, we run large sets of ΛCDM simulations and analyse them using a large variety of diagnostics, including new and powerful wavelet statistics. Among other facts, we point out (1) that dynamical evolution does not propagate discreteness noise up from the small scales at which it is introduced, and (2) that one should aim to satisfy the condition $\epsilon \sim 2d$, where ϵ is the force resolution and d is the interparticle distance. We clarify what such a condition means, and how to implement it in modern cosmological codes.

Publication: Romeo et al. (2008)

A Toomre-like Stability Criterion for the Clumpy and Turbulent Interstellar Medium — We explore the gravitational instability of clumpy and turbulent gas discs, taking into account the Larson-type scaling laws observed in giant molecular clouds (GMCs) and H I, as well as more general scaling relations. This degree of freedom is of special interest in view of the coming high-z ISM surveys, and is thus potentially important for understanding the dynamical effects of turbulence at all epochs of galaxy evolution. Our analysis shows that turbulence has a deep impact on the gravitational instability of the disc. It excites a rich variety of stability regimes, several of which have no classical counterpart. Among other diagnostics, we provide *two useful tools* for observers and simulators: (1) the stability map of turbulence, which illustrates our stability scenario and relates it to the phenomenology of interstellar turbulence: GMC/H I observations, simulations and models; (2) a Toomre-like stability criterion, $Q \geq \overline{Q}$, which applies to a large class of clumpy/turbulent discs. We make specific predictions about GMC and cold-H I turbulence, and point out the implications of our analysis for high-z galaxy surveys.

Publication: Romeo et al. (2010)

Resolving Mixing in Smoothed Particle Hydrodynamics — Standard formulations of smoothed particle hydrodynamics (SPH) are unable to resolve mixing at fluid boundaries. We use an error and stability analysis of the generalised SPH equations of motion to prove that this is due to two distinct problems. The first is a leading order error in the momentum equation. This should decrease with increasing neighbour number, but does not because numerical instabilities cause the kernel to be irregularly sampled. We identify two important instabilities: the clumping instability and the banding instability, and we show that both are cured by a suitable choice of kernel. The second problem is the local mixing instability (LMI). This occurs as particles attempt to mix on the kernel scale, but are unable to due to entropy conservation. The result is a pressure discontinuity at boundaries that pushes fluids of different entropy apart. We cure the LMI by using a weighted density estimate that ensures that pressures are single valued throughout the flow. This also gives a better volume estimate for the particles, reducing errors in the continuity and momentum equations. We demonstrate mixing in our new Optimised Smoothed Particle Hydrodynamics (OSPH) scheme using a Kelvin Helmholtz instability (KHI) test with density contrast 1:2, and the ‘blob test’ – a 1:10 density ratio gas sphere in a wind tunnel – finding excellent agreement between OSPH and Eulerian codes.

Publication: Read et al. (2010)

Systematic Uncertainties in the Determination of the Local Dark Matter Density — A precise determination of the local dark matter density and an accurate control over the corresponding uncertainties are of paramount importance for Dark Matter (DM) searches. Using very recent high-resolution numerical simulations of a Milky Way like object, we study the systematic uncertainties that affect the determination of the local dark matter density based on dynamical measurements in the Galaxy. In particular, extracting from the simulation with baryons the orientation of the Galactic stellar disk with respect to the DM distribution, we study the DM density for an observer located at ∼8 kpc from the Galactic center *on the stellar disk*, ρ_0. This quantity is found to be always larger than the average density in a spherical shell of same radius $\bar{\rho}_0$, which is the quantity inferred from dynamical measurements in the Galaxy, and to vary in the range $\rho_0/\bar{\rho}_0 = 1.01 - 1.41$. This suggests that the actual dark matter density in the solar neighbourhood is on average 21% larger than the value inferred from most dynamical measurements, and that the associated systematic errors are larger than the statistical errors recently discussed in the literature.

Publication: Pato et al. (2010)

7.2. Future Projects

Future iterations of modern codes such as RAMSES (Teyssier, 2002) and PKDGRAV (Stadel, 2001) will take full advantage of peta-scale computing resources that are becoming widely available or are planned by major supercomputing centers. This leap in computational resources will enable major progress with regard to unsolved questions like: How did our Galaxy form? What is the origin of galactic morphology — the Hubble sequence? Why does the star-formation history of the Universe decline from $z = 2$? How do galaxies acquire their baryons? Why does the baryon fraction of collapsed systems change systematically from dwarf galaxies to the Milky Way to groups and clusters of galaxies? How do galactic nuclei form and what is their relation to super-massive black holes? Exactly how do galaxies populate larger dark matter haloes? Can we reconcile the mass function of CDM haloes with the observed luminosity function of galaxies? How do bulges form — via secular evolution or minor mergers? Why do most small disk galaxies not contain a bulge? Why do clusters of galaxies contain so many more satellite galaxies than the Milky Way? These are all major questions that in principle can be studied within the framework of numerical galaxy formation. Some more specific topics that I plan to investigate are outlined below.

Star Formation in Global Simulations — The common practice of modeling star formation in galactic disks is to assume a Schmidt-law like recipe which connects the local gas density to a conversion rate into stars, using a normalization according to global averages observed in star forming galaxies (Kennicutt, 1998). While the average trends of star formation are of fundamental importance, applying these in numerical work tells us little about the underlying physical processes. Sub-kpc observations of neutral hydrogen have revealed that the average star formation relation has little correlation on a galaxy by galaxy basis (Bigiel et al., 2008), which would be challenging to predict using currently adopted recipes.

Self-consistent star formation places requirements on feedback as well. Recent work by Murray et al. (2010) demonstrates how radiative pressure from young stars dominate the feedback in large star forming clouds, hence setting, together with turbulence, the global star formation efficiency of the cloud. This results in a short star formation time-scale on order of a few million years which is supported by recent observations (Tamburro et al., 2008). To complicate things even further, the star formation efficiency is strongly dependent on metallicity, small scale gas clumpiness and the H_2 gas fraction (Gnedin et al., 2009; Krumholz et al., 2009). Without including effects of this kind, future global simulations of galaxies will keep having little predictive power. A good example of this, in a cosmological context, is from Okamoto et al. (2010) where slight changes in feedback or star formation efficiencies can turn a late type spiral galaxy into an early type.

Given the necessity for improvements, I plan to add improved algorithms of e.g. chemical networks, cooling, improved treatment of star formation and first steps towards including radiative feedback in the AMR code RAMSES. This will allow for such diverse studies as signals of primordial molecules from the dark ages (i.e. redshifts prior to recombination) to self-consistent treatment of star formation efficiencies in global simulations.

High-z Studies of Galaxy Formation, Origin of Bulges and Thick Disks — The theoretically predicted cold stream accretion (e.g. Kereš et al., 2005; Dekel & Birnboim, 2006; Kereš et al., 2009) is potentially a paradigm shift in the theory of galaxy formation, especially for large disks like our own Milky Way. To numerically resolve the cold stream accretion, and the resulting disk instabilities, can be done at high redshifts, but to reliably simulate the systems to $z = 0$ has so far proven to be highly computationally demanding. In addition, the physical processes discussed above have so far not been present.

I plan to study, at high resolution, the impact that this mode of galaxy assembly has on the angular momentum during disk formation. In addition, it is conjectured (Bournaud & Elmegreen, 2009) that the disk instabilities, i.e. the large star forming clumps, are responsible for the formation of the thick disk and the galactic bulge. Previous work has shown great promise in forming massive classical bulges this way, in contrast to the established view that these components are a result from hierarchical merging of satellites. As the thick disk and bulge of the Milky Way are important near field cosmological probes, making robust theoretical predictions for observations is a strong motivation for this research.

The Missing Baryons — Our current standard model of cosmology (ΛCDM) tells us that only $\sim 17\%$ of all matter is in baryonic form (Komatsu et al., 2009). It is natural to expect that this fraction is reflected in collapsed structures such as galaxies and clusters. Observationally, the baryon fraction is strongly dependent on the dark halo host mass; most of the baryons can be accounted for in large halos, while small haloes seem to have lost most of it. This is often referred to as called baryonic Tully-Fisher relationship (McGaugh, 2005). In our own galaxy, only one third of the expected baryons are accounted for. While less massive systems potentially can be explained by outflows from star forming regions at high redshifts or inefficient cooling due to a UV background (Hoeft et al., 2006), the Milky Way deficiency poses a problem for our understanding of galaxy formation.

In state-of-the-art simulations of Milky Way assembly, we will be able can make robust constraints on where this mass is, and in what gas phase. For example, simulations predict the existence of hot massive gas haloes around galaxies that maybe be responsible for fueling the quiescent star formation we observe today. In addition, a hot extended halo might be able to explain the origin of the Magellanic Stream, which is conjectured to arise via ram-pressure stripping against the hot halo gas. In upcoming work (Agertz et al. in prep), we demonstrate that the necessary conditions (gas densities of $n \sim 10^{-4}\,\mathrm{cm}^{-3}$) for this processes to exist, while still being in agreement with derived values form OVII absorption lines (Bregman, 2007), are naturally in place in fully cosmological simulations of Milky Way galaxies. This in turn explain a significant portion of the missing baryon problem.

Bibliography

Abadi M. G., Moore B., Bower R. G., 1999, MNRAS, 308, 947
Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003a, ApJ, 591, 499
Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003b, ApJ, 597, 21
Agertz O., Lake G., Teyssier R., Moore B., Mayer L., Romeo A. B., 2009, MNRAS, 392, 294
Agertz O., Moore B., Stadel J., Potter D., Miniati F., Read J., Mayer L., Gawryszczak A., Kravtsov A., Nordlund Å., Pearce F., Quilis V., Rudd D., Springel V., Stone J., Tasker E., Teyssier R., Wadsley J., Walder R., 2007, MNRAS, 380, 963
Agertz O., Teyssier R., Moore B., 2009, MNRAS, 397, L64
Agertz O., Teyssier R., Moore B., 2011, MNRAS, 410, 1391
Anninos W. Y., Norman M. J., 1994, ApJ, 429, 434
Balbus S. A., Hawley J. F., 1991, ApJ, 376, 214
Balsara D. S., 1995, Journal of Computational Physics, 121, 357
Barnes J., Hut P., 1986, Nature, 324, 446
Barnes J. E., Hernquist L., 1992, Nature, 360, 715
Bate M. R., Bonnell I. A., Bromm V., 2003, MNRAS, 339, 577
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Bertin G., Romeo A. B., 1988, A&A, 195, 105
Bertschinger E., 1995, ArXiv Astrophysics e-prints
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press
Birnboim Y., Dekel A., 2003, MNRAS, 345, 349
Bland-Hawthorn J., Sutherland R., Agertz O., Moore B., 2007, ApJ, 670, L109
Blitz L., Fukui Y., Kawamura A., Leroy A., Mizuno N., Rosolowsky E., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V Giant Molecular Clouds in Local Group Galaxies. pp 81–96
Bottema R., 2003, MNRAS, 344, 358
Boulanger F., Viallefond F., 1992, A&A, 266, 37
Bournaud F., Daddi E., Elmegreen B. G., Elmegreen D. M., Nesvadba N., Vanzella E., di Matteo P., Le Tiran L., Lehnert M., Elbaz D., 2008, A&A, 486, 741
Bournaud F., Duc P.-A., Emsellem E., 2008, MNRAS, 389, L8
Bournaud F., Elmegreen B. G., 2009, ApJ, 694, L158
Bournaud F., Elmegreen B. G., Elmegreen D. M., 2007, ApJ, 670, 237
Brandt A., 1977, Math. Comp, 31, 333
Bregman J. N., 2007, ARA&A, 45, 221
Brooks A. M., Governato F., Booth C. M., Willman B., Gardner J. P., Wadsley J., Stinson G., Quinn T., 2007, ApJ, 655, L17
Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396
Bryan G. L., Norman M. L., 1997, in Clarke D. A., West M. J., eds, ASP Conf. Ser. 123: Computational Astrophysics; 12th Kingston Meeting on Theoretical Astrophysics Simulating X-Ray Clusters with Adaptive Mesh Refinement. pp 363–+
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Bryan G. L., Norman M. L., Stone J. M., Cen R., Ostriker J. P., 1995, Comp. Phys. Comm., 89, 149
Burkert A., 2006, Comptes Rendus Physique, 7, 433
Carollo D., Beers T. C., Lee Y. S., Chiba M., Norris J. E., Wilhelm R., Sivarani T., Marsteller B., Munn J. A., Bailer-Jones C. A. L., Fiorentin P. R., York D. G., 2007, Nature, 450, 1020
Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499

Ceverino D., Klypin A., 2007, preprint (arXiv:0712.3285)
Chandrasekhar S., 1961, Hydrodynamic and hydromagnetic stability. International Series of Monographs on Physics, Oxford: Clarendon, 1961
Chung T. J., 2002, Computational Fluid Dynamics. Cambridge University Press
Colella P., Glaz H. M., 1985, JCPh, 59, 264
Colella P., Woodward P. R., 1984, JCPh, 54, 174
Corbelli E., 2003, MNRAS, 342, 199
Courteau S., 1997, AJ, 114, 2402
Courteau S., de Jong R. S., Broeils A. H., 1996, ApJ, 457, L73+
Cowie L. L., Hu E. M., Songaila A., 1995, AJ, 110, 1576
Croft R. A. C., Di Matteo T., Springel V., Hernquist L., 2009, MNRAS, 400, 43
Daddi E., Cimatti A., Renzini A., Fontana A., Mignoli M., Pozzetti L., Tozzi P., Zamorani G., 2004, ApJ, 617, 746
Dalcanton J. J., Shectman S. A., 1996, ApJ, 465, L9+
de Avillez M. A., Breitschwerdt D., 2004, A&A, 425, 899
de Avillez M. A., Breitschwerdt D., 2005, A&A, 436, 585
Debattista V. P., Mayer L., Carollo C. M., Moore B., Wadsley J., Quinn T., 2006, ApJ, 645, 209
Dehnen W., Binney J., 1998, MNRAS, 294, 429
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Dekel A., Birnboim Y., Engel G., Freundlich J., Goerdt T., Mumcuoglu M., Neistein E., Pichon C., Teyssier R., Zinger E., 2009, Nature, 457, 451
Dekel A., Sari R., Ceverino D., 2009, ArXiv e-prints
Dekel A., Silk J., 1986, ApJ, 303, 39
Deul E. R., van der Hulst J. M., 1987, A&AS, 67, 509
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Dib S., Bell E., Burkert A., 2006, ApJ, 638, 797
Dickey J. M., Brinks E., 1993, ApJ, 405, 153
Dickey J. M., Hanson M. M., Helou G., 1990, ApJ, 352, 522
Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J., 2008, Nature, 454, 735
Dijkstra M., Haiman Z., Rees M. J., Weinberg D. H., 2004, ApJ, 601, 666
Dolag K., Vazza F., Brunetti G., Tormen G., 2005, MNRAS, 364, 753
Doroshkevich A. G., Zeldovich I. B., 1981, Zhurnal Eksperimental noi i Teoreticheskoi Fiziki, 80, 801
Dubois Y., Teyssier R., 2008, A&A, 477, 79
Dutton A. A., Conroy C., van den Bosch F. C., Prada F., More S., 2010, MNRAS, 407, 2
Dutton A. A., van den Bosch F. C., 2009, MNRAS, 396, 141
Efstathiou G., 2000, MNRAS, 317, 697
Einasto J., 1969, Astrofizika, 5, 137
Eisenstein D. J., Hu W., 1999, ApJ, 511, 5
Eisenstein D. J., Hut P., 1998, ApJ, 498, 137
Elmegreen B. G., Bournaud F., Elmegreen D. M., 2008, ApJ, 688, 67
Elmegreen B. G., Elmegreen D. M., 2005, ApJ, 627, 632
Elmegreen B. G., Elmegreen D. M., 2006, ApJ, 650, 644
Elmegreen B. G., Elmegreen D. M., Fernandez M. X., Lemonias J. J., 2009, ApJ, 692, 12
Elmegreen B. G., Kaufman M., Thomasson M., 1993, ApJ, 412, 90
Elmegreen B. G., Scalo J., 2004, ARA&A, 42, 211
Elmegreen D. M., Elmegreen B. G., Hirst A. C., 2004, ApJ, 604, L21
Elmegreen D. M., Elmegreen B. G., Ravindranath S., Coe D. A., 2007, ApJ, 658, 763
Engargiola G., Plambeck R. L., Rosolowsky E., Blitz L., 2003, ApJS, 149, 343
Escala A., Larson R. B., 2008, ApJ, 685, L31
Eskridge P. B., Frogel J. A., Pogge R. W., Quillen A. C., Berlind A. A., Davies R. L., DePoy D. L., Gilbert K. M., Houdashelt M. L., Kuchinski L. E., Ramírez S. V., Sellgren K., Stutz A., Terndrup D. M., Tiede G. P., 2002, ApJS, 143, 73
Faber T. E., 1995, Fluid dynamics for physicists. Cambridge, New York: Cambridge University Press, |c1995
Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189

Ferguson A. M. N., Gallagher J. S., Wyse R. F. G., 1998, AJ, 116, 673
Ferrière K. M., 2001, Reviews of Modern Physics, 73, 1031
Fleck Jr. R. C., 1981, ApJ, 246, L151
Förster Schreiber N. M., Genzel R., Lehnert M. D., Bouché N., Verma A., Erb D. K., Shapley A. E., Steidel C. C., Davies R., Lutz D., 2006, ApJ, 645, 1062
Frenk C. S., et al 1999, ApJ, 525, 554
Fryxell B., Olson K., Ricker P., Timmes F. X., Zingale M., Lamb D. Q., MacNeice P., Rosner R., Truran J. W., Tufo H., 2000, ApJS, 131, 273
Fuchs B., 2001, A&A, 368, 107
Fukunaga M., Tosa M., 1989, PASJ, 41, 241
Gammie C. F., 2001, ApJ, 553, 174
Gammie C. F., Ostriker J. P., Jog C. J., 1991, ApJ, 378, 565
Garnett D. R., 2002, ApJ, 581, 1019
Genzel R., Burkert A., Bouché N., Cresci G., Förster Schreiber N. M., Shapley A., Shapiro K., Tacconi L. J., Buschkamp P., Cimatti A., Daddi E., 2008, ApJ, 687, 59
Genzel R., Tacconi L. J., Eisenhauer F., Förster Schreiber N. M., Cimatti A., Daddi E., Bouché N., Davies R., Lehnert M. D., Lutz D., Nesvadba N., Verma A., 2006, Nature, 442, 786
Gerritsen J. P. E., 1997, PhD thesis, , Groningen University, the Netherlands, (1997)
Gerritsen J. P. E., Icke V., 1997, A&A, 325, 972
Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375
Gnedin N. Y., 2000, ApJ, 535, 530
Gnedin N. Y., Kravtsov A. V., 2010a, preprint (ArXiv: 1004.0003)
Gnedin N. Y., Kravtsov A. V., 2010b, ApJ, 714, 287
Gnedin N. Y., Tassis K., Kravtsov A. V., 2009, ApJ, 697, 55
Gnedin O. Y., Weinberg D. H., Pizagno J., Prada F., Rix H., 2007, ApJ, 671, 1115
Godunov S., 1959, Math. Sbornik, 47, 271
Goldreich P., Lynden-Bell D., 1965a, MNRAS, 130, 97
Goldreich P., Lynden-Bell D., 1965b, MNRAS, 130, 125
Governato F., Brook C., Mayer L., Brooks A., Rhee G., Wadsley J., Jonsson P., Willman B., Stinson G., Quinn T., Madau P., 2010, Nature, 463, 203
Governato F., Brook C. B., Brooks A. M., Mayer L., Willman B., Jonsson P., Stilp A. M., Pope L., Christensen C., Wadsley J., Quinn T., 2009, MNRAS, 398, 312
Governato F., Mayer L., Wadsley J., Gardner J. P., Willman B., Hayashi E., Quinn T., Stadel J., Lake G., 2004, ApJ, 607, 688
Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, MNRAS, 374, 1479
Graham A. W., Worley C. C., 2008, MNRAS, 388, 1708
Gregori G., Miniati F., Ryu D., Jones T. W., 2000, ApJ, 543, 775
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Haardt F., Madau P., 1996, ApJ, 461, 20
Hansen S. H., Agertz O., Joyce M., Stadel J., Moore B., Potter D., 2007, ApJ, 656, 631
Heath D. J., 1977, MNRAS, 179, 351
Hennebelle P., Audit E., 2007, A&A, 465, 431
Hernquist L., Katz N., 1989, ApJS, 70, 419
Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles. New York: McGraw-Hill, 1981
Hoeft M., Yepes G., Gottlöber S., Springel V., 2006, MNRAS, 371, 401
Inogamov N. A., 1999, Astrophysics and Space Physics Reviews, 10, 1
Jog C. J., 1992, ApJ, 390, 378
Jog C. J., Ostriker J. P., 1988, ApJ, 328, 404
Jog C. J., Solomon P. M., 1984, ApJ, 276, 114
Jogee S., Miller S., Penner K., Bell E. F., Conselice C., Skelton R. E., Somerville R. S., Rix H.-W., Barazza F. D., Barden M., Borch A., Beckwith S. V., Caldwell J. A. R., Häussler B., 2008, in Funes J. G., Corsini E. M., eds, Astronomical Society of the Pacific Conference Series Vol. 396 of Astronomical Society of the Pacific Conference Series, Frequency and Impact of Galaxy Mergers and Interactions over the Last 7 Gyr.

pp 337–+

Jones T. W., Ryu D., Tregillis I. L., 1996, ApJ, 473, 365

Joung M. K. R., Mac Low M.-M., 2006, ApJ, 653, 1266

Julian W. H., Toomre A., 1966, ApJ, 146, 810

Kalirai J. S., Hansen B. M. S., Kelson D. D., Reitzel D. B., Rich R. M., Richer H. B., 2008, The Astrophysical Journal, 676, 594

Kamphuis J., Sancisi R., 1993, A&A, 273, L31+

Katz N., 1992, ApJ, 391, 502

Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19

Kay S. T., Pearce F. R., Frenk C. S., Jenkins A., 2002, MNRAS, 330, 113

Kennicutt Jr. R. C., 1998, ApJ, 498, 541

Kennicutt Jr. R. C., Calzetti D., Walter F., et al. 2007, ApJ, 671, 333

Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160

Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2

Kessel-Deynet O., Burkert A., 2003, MNRAS, 338, 545

Khokhlov A. M., 1998, JCPh, 143, 519

Kim J., Balsara D., Mac Low M.-M., 2001, Journal of Korean Astronomical Society, 34, 333

Kim W.-T., Ostriker E. C., 2001, ApJ, 559, 70

Kim W.-T., Ostriker E. C., 2007, ApJ, 660, 1232

Kim W.-T., Ostriker E. C., Stone J. M., 2002, ApJ, 581, 1080

Kim W.-T., Ostriker E. C., Stone J. M., 2003, ApJ, 599, 1157

Klein R. I., McKee C. F., Colella P., 1994, ApJ, 420, 213

Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82

Klypin A., Zhao H., Somerville R. S., 2002, ApJ, 573, 597

Knop R. A., Aldering G., Amanullah R., Astier P., et al. 2003, ApJ, 598, 102

Komatsu E., Dunkley J., Nolta M. R., Bennett C. L., Gold B., Hinshaw G., Jarosik N., Larson D., Limon M., Page L., Spergel D. N., Halpern M., Hill R. S., Kogut A., Meyer S. S., Tucker G. S., Weiland J. L., Wollack E., Wright E. L., 2009, ApJS, 180, 330

Kormendy J., Kennicutt Jr. R. C., 2004, ARA&A, 42, 603

Kravtsov A. V., 1999, PhD thesis, New Mexico State University

Kravtsov A. V., 2003, ApJ, 590, L1

Kravtsov A. V., Klypin A., Hoffman Y., 2002, ApJ, 571, 563

Kroupa P., Tout C. A., Gilmore G., 1993, MNRAS, 262, 545

Krumholz M. R., McKee C. F., Tumlinson J., 2009, ApJ, 699, 850

Krumholz M. R., Tan J. C., 2007, ApJ, 654, 304

Laney C., 1998, Computational gasdynamics. CUP, Cambridge

Larson R. B., 1981, MNRAS, 194, 809

Lattanzio J. C., Monaghan J. J., Pongracic H., Schwarz M. P., 1986, j-SIAM-J-SCI-STAT-COMP, 7, 591

Laurikainen E., Salo H., Buta R., Knapen J. H., Comerón S., 2010, MNRAS, 405, 1089

Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782

Leveque R. J., 1998, in Steiner O., Gautschy A., eds, Saas-Fee Advanced Course 27: Computational Methods for Astrophysical Fluid Flow. Nonlinear Conservation Laws and Finite Volume Methods. p. 1

Li C., White S. D. M., 2009, MNRAS, 398, 2177

Lifshitz E. M., 1946, J. Phys. USSR, 10, 116

Lin D. N. C., Pringle J. E., 1987, MNRAS, 225, 607

Lucy L. B., 1977, AJ, 82, 1013

Mac Low M., Ferrara A., 1999, ApJ, 513, 142

Mac Low M.-M., Balsara D. S., Kim J., de Avillez M. A., 2005, ApJ, 626, 864

Mac Low M.-M., Klessen R. S., 2004, Reviews of Modern Physics, 76, 125

Mac Low M.-M., McKee C. F., Klein R. I., Stone J. M., Norman M. L., 1994, ApJ, 433, 757

Mac Low M.-M., Zahnle K., 1994, ApJ, 434, L33

Macciò A. V., Moore B., Stadel J., 2006, ApJ, 636, L25

Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715

Marcolini A., D’Ercole A., Brighenti F., Recchi S., 2006, MNRAS, 371, 643
Marri S., White S. D. M., 2003, MNRAS, 345, 561
Martig M., Bournaud F., 2010, ApJ, 714, L275
Mathewson D. S., Ford V. L., Buchhorn M., 1992, ApJS, 81, 413
Mathieu A., Merrifield M. R., Kuijken K., 2002, MNRAS, 330, 251
Mayer L., Quinn T., Wadsley J., Stadel J., 2002, Science, 298, 1756
McGaugh S. S., 2005, ApJ, 632, 859
McGaugh S. S., Schombert J. M., Bothun G. D., de Blok W. J. G., 2000, ApJ, 533, L99
McGaugh S. S., Schombert J. M., de Blok W. J. G., Zagursky M. J., 2010, ApJ, 708, L14
McKee C. F., Ostriker E. C., 2007, ARA&A, 45, 565
McKee C. F., Ostriker J. P., 1977a, ApJ, 218, 148
McKee C. F., Ostriker J. P., 1977b, ApJ, 218, 148
Meurer G. R., Carignan C., Beaulieu S. F., Freeman K. C., 1996, AJ, 111, 1551
Miniati F., Colella P. J., 2006, Comp. Phys. submitted, e-print:astro-ph/0608156
Miniati F., Jones T. W., Ryu D., 1999, ApJ, 517, 242
Misiriotis A., Xilouris E. M., Papamastorakis J., Boumis P., Goudis C. D., 2006, A&A, 459, 113
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Monaghan J. J., 1992, ARA&A, 30, 543
Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19
More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2010, ArXiv e-prints
Mori M., Burkert A., 2000, ApJ, 538, 559
Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Murray N., Quataert E., Thompson T. A., 2010, ApJ, 709, 191
Murray S. D., White S. D. M., Blondin J. M., Lin D. N. C., 1993, ApJ, 407, 588
Nair P. B., Abraham R. G., 2010, ApJS, 186, 427
Navarro J. F., Benz W., 1991, ApJ, 380, 320
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Navarro J. F., Steinmetz M., 2000, ApJ, 538, 477
Navarro J. F., White S. D. M., 1993, MNRAS, 265, 271
Navarro J. F., White S. D. M., 1994, MNRAS, 267, 401
Neistein E., van den Bosch F. C., Dekel A., 2006, MNRAS, 372, 933
Nulsen P. E. J., 1982, MNRAS, 198, 1007
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Okamoto T., Eke V. R., Frenk C. S., Jenkins A., 2005, MNRAS, 363, 1299
Okamoto T., Frenk C. S., Jenkins A., Theuns T., 2010, MNRAS, 406, 208
Okamoto T., Jenkins A., Eke V. R., Quilis V., Frenk C. S., 2003, MNRAS, 345, 429
Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265
Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, MNRAS, 406, 2325
O’Shea B. W., Bryan G., Bordner J., Norman M. L., Abel T., Harkness R., Kritsuk A., 2004, ArXiv Astrophysics e-prints
O’Shea B. W., Nagamine K., Springel V., Hernquist L., Norman M. L., 2005, ApJS, 160, 1
Overzier R. A., Heckman T. M., Kauffmann G., Seibert M., Rich R. M., Basu-Zych A., Lotz J., Aloisi A., Charlot S., Hoopes C., Martin D. C., Schiminovich D., Madore B., 2008, ApJ, 677, 37
Padmanabhan T., 1993, Structure Formation in the Universe. Cambridge University Press
Pato M., Agertz O., Bertone G., Moore B., Teyssier R., 2010, Phys. Rev. D, 82, 023531
Pearce F. R., Jenkins A., Frenk C. S., Colberg J. M., White S. D. M., Thomas P. A., Couchman H. M. P., Peacock J. A., Efstathiou G., The Virgo Consortium 1999, ApJ, 521, L99
Peebles P. J. E., 1969, ApJ, 155, 393
Pelupessy F. I., Papadopoulos P. P., 2009, ApJ, 707, 954
Petric A. O., Rupen M. P., 2007, AJ, 134, 1952
Pfalzner S., Gibbon P., 1996, Many-Body Tree Methods in Physics. Cambridge University Press

Piontek F., Steinmetz M., 2009a, preprint (ArXiv: 0909.4156)
Piontek F., Steinmetz M., 2009b, preprint (ArXiv: 0909.4167)
Piontek R. A., Ostriker E. C., 2004, ApJ, 601, 905
Piontek R. A., Ostriker E. C., 2005, ApJ, 629, 849
Pizagno J., Prada F., Weinberg D. H., Rix H., Pogge R. W., Grebel E. K., Harbeck D., Blanton M., Brinkmann J., Gunn J. E., 2007, AJ, 134, 945
Pohlen M., Trujillo I., 2006, A&A, 454, 759
Press W. H., Schechter P., 1974, ApJ, 187, 425
Price D., 2005, ArXiv Astrophysics e-prints
Price D. J., 2008, Journal of Computational Physics, 227, 10040
Puche D., Westpfahl D., Brinks E., Roy J.-R., 1992, AJ, 103, 1841
Quilis V., Moore B., Bower R., 2000, Science, 288, 1617
Quinn T., Katz N., Efstathiou G., 1996, MNRAS, 278, L49
Rafelski M., Wolfe A. M., Cooke J., Chen H., Armandroff T. E., Wirth G. D., 2009, ApJ, 703, 2033
Rafikov R. R., 2001, MNRAS, 323, 445
Raiteri C. M., Villata M., Navarro J. F., 1996, A&A, 315, 105
Rasera Y., Teyssier R., 2006, A&A, 445, 1
Read J. I., Hayfield T., Agertz O., 2010, MNRAS, 405, 1513
Read J. I., Lake G., Agertz O., Debattista V. P., 2008, MNRAS, 389, 1041
Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, MNRAS, 374, 2
Regan J. A., Haehnelt M. G., Viel M., 2007, MNRAS, 374, 196
Rhode K. L., Salzer J. J., Westpfahl D. J., Radice L. A., 1999, AJ, 118, 323
Ritchie B. W., Thomas P. A., 2001, MNRAS, 323, 743
Robertson B., Yoshida N., Springel V., Hernquist L., 2004, ApJ, 606, 32
Robertson B. E., Bullock J. S., 2008, ApJ, 685, L27
Robertson B. E., Kravtsov A. V., 2008, ApJ, 680, 1083
Romeo A. B., 1992, MNRAS, 256, 307
Romeo A. B., 1994, A&A, 286, 799
Romeo A. B., Agertz O., Moore B., Stadel J., 2008, ApJ, 686, 1
Romeo A. B., Burkert A., Agertz O., 2010, MNRAS, 407, 1223
Rosen A., Bregman J. N., 1995, ApJ, 440, 634
Rosolowsky E., Keto E., Matsushita S., Willner S. P., 2007, ApJ, 661, 830
Roškar R., Debattista V. P., Stinson G. S., Quinn T. R., Kaufmann T., Wadsley J., 2008, ApJ, 675, L65
Sawala T., Scannapieco C., Maio U., White S., 2010, MNRAS, 402, 1599
Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2006, MNRAS, 371, 1125
Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2008, MNRAS, 389, 1137
Scannapieco C., White S. D. M., Springel V., Tissera P. B., 2009, MNRAS, 396, 696
Schmidt M., 1959, ApJ, 129, 243
Schuster K. F., Kramer C., Hitschfeld M., Garcia-Burillo S., Mookerjea B., 2007, A&A, 461, 143
Seigar M. S., Barth A. J., Bullock J. S., 2008, MNRAS, 389, 1911
Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437
Sellwood J. A., Balbus S. A., 1999, ApJ, 511, 660
Shapiro K. L., Genzel R., Förster Schreiber N. M., Tacconi L. J., Bouché N., Cresci G., Davies R., Eisenhauer F., Johansson P. H., Krajnović D., Lutz D., Naab T., Arimoto N., Arribas S., 2008, ApJ, 682, 231
Shapiro S. L., Teukolsky S. A., 1983, Black holes, white dwarfs, and neutron stars: The physics of compact objects. New York, Wiley-Interscience
Shen J., Sellwood J. A., 2006, MNRAS, 370, 2
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Shu F. H., 1992, Physics of Astrophysics, Vol. II. Physics of Astrophysics, Vol. II, by Frank H. Shu. Published by University Science Books, ISBN 0-935702-65-2, 476pp, 1992.
Slyz A. D., Devriendt J. E. G., Bryan G., Silk J., 2005, MNRAS, 356, 737
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Sommer-Larsen J., Götz M., Portinari L., 2003, ApJ, 596, 47
Springel V., 2005, MNRAS, 364, 1105

Springel V., Hernquist L., 2002, MNRAS, 333, 649
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Yoshida N., White S. D. M., 2001, New Astronomy, 6, 79
Stadel J., Potter D., Moore B., Diemand J., Madau P., Zemp M., Kuhlen M., Quilis V., 2009, MNRAS, 398, L21
Stadel J. G., 2001, Ph.D. Thesis
Stark A. A., Brand J., 1989, ApJ, 339, 763
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Stinson G. S., Dalcanton J. J., Quinn T., Gogarten S. M., Kaufmann T., Wadsley J., 2009, MNRAS, 395, 1455
Stone J. M., Norman M. L., 1992a, ApJS, 80, 753
Stone J. M., Norman M. L., 1992b, ApJS, 80, 791
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Tamburro D., Rix H., Walter F., Brinks E., de Blok W. J. G., Kennicutt R. C., Mac Low M., 2008, AJ, 136, 2872
Taniguchi Y., Shioya Y., 2001, ApJ, 547, 146
Tasker E. J., Bryan G. L., 2006, ApJ, 641, 878
Tasker E. J., Tan J. C., 2009, ApJ, 700, 358
Tassis K., Kravtsov A. V., Gnedin N. Y., 2008, ApJ, 672, 888
Teyssier R., 2002, A&A, 385, 337
Teyssier R., Pires S., Prunet S., Aubert D., Pichon C., Amara A., Benabed K., Colombi S., Refregier A., Starck J., 2009, A&A, 497, 335
Thacker R. J., Tittley E. R., Pearce F. R., Couchman H. M. P., Thomas P. A., 2000, MNRAS, 319, 619
Thielemann F., Nomoto K., Yokoi K., 1986, A&A, 158, 17
Thoul A. A., Weinberg D. H., 1996, ApJ, 465, 608
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tittley E. R., Pearce F. R., Couchman H. M. P., 2001, ApJ, 561, 69
Toomre A., 1964, ApJ, 139, 1217
Toomre A., 1981, in Fall S. M., Lynden-Bell D., eds, Structure and Evolution of Normal Galaxies What amplifies the spirals. pp 111–136
Toomre A., Toomre J., 1972, ApJ, 178, 623
Toro E. F., 1999, Riemann Solvers and Numerical Methods for Fluid Dynamics, 2nd edition. Springer-Verlag, Berlin/Heidelberg/New York
Truelove J. K., Klein R. I., McKee C. F., Holliman II J. H., Howell L. H., Greenough J. A., 1997, ApJ, 489, L179+
Tully R. B., Fisher J. R., 1977, A&A, 54, 661
Valcke S., de Rijcke S., Dejonghe H., 2008, MNRAS, 389, 1111
van Albada T. S., Bahcall J. N., Begeman K., Sancisi R., 1985, ApJ, 295, 305
van den Bergh S., McClure R. D., 1994, ApJ, 425, 205
van der Hulst J. M., 1996, in Skillman E. D., ed., The Minnesota Lectures on Extragalactic Neutral Hydrogen Vol. 106 of Astronomical Society of the Pacific Conference Series, Bubbles and Holes in the Interstellar Medium. pp 47–+
van Leer B., 1977, JCPh, 23, 276
van Leer B., 1979, JCPh, 32, 101
Vietri M., Ferrara A., Miniati F., 1997, ApJ, 483, 262
Vikhlinin A., Markevitch M., Murray S. S., 2001, ApJ, 549, L47
Wada K., Meurer G., Norman C. A., 2002, ApJ, 577, 197
Wada K., Norman C. A., 2007, ApJ, 660, 276
Wadsley J. W., Stadel J., Quinn T., 2004, New Astronomy, 9, 137
Wang Q. D., Yao Y., Tripp T. M., Fang T., Cui W., Nicastro F., Mathur S., Williams R. J., Song L., Croft R., 2005, ApJ, 635, 386
Watkins S. J., Bhattal A. S., Francis N., Turner J. A., Whitworth A. P., 1996, A&AS, 119, 177
Weinberg S., 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity. Wiley-VCH

Weinzirl T., Jogee S., Khochfar S., Burkert A., Kormendy J., 2009, ApJ, 696, 411
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Wolfe A. M., Chen H., 2006, ApJ, 652, 981
Wong T., Blitz L., 2002, ApJ, 569, 157
Xue X. X., Rix H. W., Zhao G., Re Fiorentin P., Naab T., Steinmetz M., van den Bosch F. C., Beers T. C., Lee Y. S., Bell E. F., Rockosi C., Yanny B., Newberg H., Wilhelm R., Kang X., Smith M. C., Schneider D. P., 2008, ApJ, 684, 1143
Zel'Dovich Y. B., 1970, A&A, 5, 84
Zhang W., Li C., Kauffmann G., Zou H., Catinella B., Shen S., Guo Q., Chang R., 2009, MNRAS, 397, 1243
Zheng W., Mikles V. J., Mainieri V., Hasinger G., Rosati P., Wolf C., Norman C., Szokoly G., Gilli R., Tozzi P., Wang J. X., Zirm A., Giacconi R., 2004, ApJS, 155, 73

Acknowledgements

I am very grateful to my supervisor Ben Moore for giving me the opportunity of working on a large number of interesting topics in astrophysics, especially galaxy formation, and for creating the relaxed, inspirational and fun working environment that exists in the institute.

I am indebted to Romain Teyssier for introducing me to the treacherous world of hydrodynamics and for letting me join the RAMSES crew. Thank you Romain for always taking the time to chat about ongoing projects; I have benefited immensely from our countless number of discussion. Thanks to both Ben and Romain for teaching me how to (mis)use important scientific words and phrases such as: *unrivaled, cutting-edge, fully resolved, fundamental, state-of-the-art, beyond the scope of, understood, unparalleled, numerically converged, turbulent* etc.

Thank you Alessandro Romeo for introducing me to real science, being a great teacher and friend, and for connecting me to the scientific community!

My life here would not have been the same if I hadn't met all the great people at the Uni! Thanks to my friends Doug, Joachim, Justin, Jonathan, Lea, Simone, Suzanne, Hanni, Tristen, Francis, Aaron (up and downstairs!), Karen, George, Sinziana, Pedro, all Tobiases, and everybody else I've forgotten to mention but had an awesome time with.

On a personal note, I want to thank my wife Ann-Sofie for always being there for me: *Utan dig känns allting betydelselöst!* I am of course also forever grateful to my family: *Mamma, Pappa och Henrik, utan er skulle jag aldrig varit där jag är idag.*

Curriculum Vitae

Bibliographical Information

Name:	Oscar Lennart Agertz
Date of Birth:	August 31, 1980
Place of Birth:	Göteborg, Sweden
Citizenship:	Swedish

Research Interests

Numerical studies of structure and galaxy formation and evolution
Structure and dynamics of the interstellar medium
Computational astrophysics and hydrodynamics
Star formation in galactic disks

Education

2005-2010	Institute for Theoretical Physics University of Zürich, Switzerland Ph.D. in Physics Thesis Title: *Light in the Heart of Darkness: Galaxy Formation from a Numerical Perspective* Advisor: Prof. Ben Moore
1999-2004	Department of Physics University of Gothenburg, Sweden M.Sci. in Physics Thesis Title: *Structure Formation in the Early Universe* Advisor: Dr. Alessandro B. Romeo

Teaching Experience

2009	Assistant in *Computational Astrophysics*, Lecturer: Prof. Romain Teyssier
2007/2008	Assistant in *Classical Mechanics*, Lecturer: Prof. Ben Moore
2006/2007	Assistant in *Introductory Astrophysics*, Lecturer: Prof. Ben Moore, Kim-Vy Tran and Lucio Mayer
2006/2007	Assistant in *Classical Mechanics*, Lecturer: Prof. Ben Moore
2005	Assistant in *Astrophysics & Cosmology I*, Lecturer: Prof. Ben Moore

Student Supervision

2010	Institute for Theoretical Physics University of Zürich, Switzerland Simon Grimm Semester Thesis: *The Milky Way-Andromeda Merger*

Schools

Mar. 2007	Mürren, Switzerland *The Origin of the Galaxy and Local Group* 37th Saas-Fee Advanced Course of the Swiss Society for Astrophysics and Astronomy

Conferences

Apr. 2010	University of Florida, Gainesville, USA “From Stars to Galaxies - Connecting our Understanding of Star and Galaxy Formation” Talk: *Formation of Extended Galactic Disks in Cosmological AMR Simulations*
Dec. 2009	University of Sydney, Sydney, Australia “Missing Baryons, Accretion and Feedback in Galaxies” Talk: *Disc formation and the origin of clumpy galaxies at high redshift*
Aug. 2009	University of Santa Cruz, Santa Cruz, USA “Galaxy Formation Workshop" Talk: *Disc formation and the origin of clumpy galaxies at high redshift*
July 2008	Ascona, Switzerland “Frontier in Computational Astrophysics” Talk: *Issues with numerical hydrodynamics*
Jun. 2008	Copenhagen, Denmark “IAU Symposium 254” Poster: *Turbulence in galactic disks: The role of self-gravity and supernova feedback*
Sep. 2007	Zürich, Switzerland “From Protostellar Cores to Disk Galaxies” Talk: *Fundamental differences between SPH and grid methods*
May 2007	Chamonix, France “Structure Formation in the Universe” Poster: *Fundamental differences between SPH and grid methods*
Jun. 2006	Munich, Germany Talk: *Simulations of two-phase media*

Publication List

Refereed Publications

1. S. H. Hansen, **O. Agertz**, M. Joyce, J. Stadel, B.Moore, D. Potter
   An Alternative to Grids and Glasses: Quaquaversal Pre-Initial Conditions for N-Body Simulations
   2007, ApJ, 656, 631
2. **O. Agertz**, B. Moore, J. Stadel et al.
   Fundamental differences between SPH and grid methods
   2007, MNRAS, 380, 963
3. J. Bland-Hawthorn, R. Sutherland, **O. Agertz**, B. Moore
   The Source of Ionization along the Magellanic Stream
   2007, ApJ, 670, L109
4. J. I. Read, G. Lake, **O. Agertz**, V. P. Debattista
   Thin, thick and dark discs in Λ*CDM*
   2008, MNRAS, 389, 104
5. A. B. Romeo, **O. Agertz**, B. Moore, J. Stadel
   Discreteness Effects in Λ*CDM Simulations: A Wavelet-Statistical View*
   2008, ApJ, 686, 1
6. **O. Agertz**, G. Lake, R. Teyssier, B. Moore, L. Mayer, A.B. Romeo
   Large-scale galactic turbulence: can self-gravity drive the observed HI velocity dispersions?
   2009, MNRAS, 392, 294
7. **O. Agertz**, R. Teyssier, B. Moore
   Disc formation and the origin of clumpy galaxies at high redshift
   2009, MNRAS, 397, L64
8. A. B. Romeo, A. Burkert, **O. Agertz**
   A Toomre-like stability criterion for the clumpy and turbulent interstellar medium
   2010, MNRAS, 904R
9. J. I. Read, T. Hayfield, **O. Agertz**
   Resolving mixing in smoothed particle hydrodynamics
   2010, MNRAS, 405, 1513
10. M. Pato, **O. Agertz**. G. Bertone, B. Moore, R. Teyssier
    Systematic uncertainties in the determination of the local dark matter density
    2010, Ph.RvD, 82, 2
11. **O. Agertz**, R. Teyssier, B. Moore
    The formation of disk galaxies in a LCDM universe
    2010, MNRAS accepted

Conference Proceedings

1. J. I. Read, G. Lake, **O. Agertz**, V. P. Debattista
   A dark disc in the Milky Way
   2008, Astronomische Nachrichten, Volume 329, Issue , p.1022
2. J. I. Read, V. P. Debattista, **O. Agertz**, L. Mayer, A. M. Brooks, F. Governato, G. Lake
   A dark matter disc in the Milky Way
   Proceedings of "Identification of dark matter 2008". August 18-22, 2008, Stockholm, Sweden, p.48
3. **O. Agertz**, G. Lake, B. Moore, L. Mayer. R. Teyssier, A. B. Romeo
   Turbulence in galactic disks: The role of self-gravity and supernova feedback
   IAU Symposium, Volume 254, Cambridge University Press, 2009, p.1
4. A. B. Romeo, **O. Agertz**, B. Moore, J. Stadel
   Discreteness Effects in Λ*CDM Simulations*
   IAU Symposium, Volume 254, Cambridge University Press, 2009, p.63